

Universidade de Lisboa
Faculdade de Ciências
Departamento de Estatística e Investigação Operacional



POPULAÇÃO PORTUGUESA POR NÍVEL DE ESCOLARIDADE,
SEXO E GRUPO ETÁRIO: UMA ANÁLISE PROSPETIVA PARA
2001-2031

por

Inês Raquel Campos Rodrigues

Trabalho de Projeto

Mestrado em Bioestatística

2012

Universidade de Lisboa
Faculdade de Ciências
Departamento de Estatística e Investigação Operacional



POPULAÇÃO PORTUGUESA POR NÍVEL DE ESCOLARIDADE,
SEXO E GRUPO ETÁRIO: UMA ANÁLISE PROSPETIVA PARA
2001-2031

por

Inês Raquel Campos Rodrigues

Trabalho de Projeto orientado pela Prof.^a Doutora Maria Isabel Calisto Frade Barão
e coorientado pela Prof.^a Doutora Maria do Rosário Oliveira Martins

Mestrado em Bioestatística

2012

Resumo

O acesso a informação de qualidade relativa à estrutura populacional num dado momento, à sua evolução nos anos recentes e à forma como a dinâmica demográfica poderá variar no futuro desempenha um papel fundamental na tomada de decisão em variados sectores. O nível de instrução é um dos atributos que levam a maior distinção entre os membros de uma população; neste contexto, o objetivo central deste trabalho consistiu na elaboração de projeções demográficas para a população portuguesa por sexo, grupos etários quinquenais e nível de escolaridade, para o período 2001-2031. Recorreu-se, para tal, ao modelo de projeção multiestado. Um modelo multiestado é definido como um modelo para um processo estocástico, que em cada instante ocupa um de entre um conjunto de estados discretos, definidos pelo conjunto de características consideradas. Para além de estimar as probabilidades de transição entre níveis de escolaridade, foi necessário considerar os diferenciais relativos à fecundidade, mortalidade e migrações, por grau de instrução. O cálculo dos efetivos populacionais teve por base o método de componentes por coorte multiestado, traduzido no recurso a uma matriz de Leslie por blocos. Foram considerados dois cenários de evolução, um em que os níveis de escolarização verificados entre 2001 e 2006 se mantêm constantes no restante período e outro em que os níveis de escolarização seguirão a tendência verificada na última década. Neste segundo cenário, foi modelada a tendência das proporções de estado para cada sexo e grupo etário, recorrendo a modelos de razão contínua.

Os resultados apontam para o aumento da proporção de indivíduos que completam níveis de escolaridade mais elevados, em todos os grupos etários de ambos os sexos. De acordo com o cenário de tendência, a percentagem de indivíduos com 65 ou mais anos com um curso de Ensino Superior, que se situava em 2,4% em 2001, deverá aumentar até 11,3% em 2031. Entre as mulheres, apenas 7,3% tinha concluído algum grau de Ensino Superior em 2001, valor que deverá subir para 22% em 2031; já entre os homens, este valor era de apenas 5,7% no ano de partida e deverá subir até 2031, atingindo os 13,2%. Dado que a passagem a um nível de escolaridade mais elevado é um evento irreversível, podemos esperar que a proporção de idosos com níveis de instrução superiores continue a aumentar, já que a escolarização das coortes mais jovens parece continuar a evoluir positivamente.

Pretende-se que este trabalho sirva de base ao estudo da forma como o envelhecimento demográfico e o aumento dos níveis de escolarização poderão contribuir para o planeamento e monitorização de políticas públicas, nomeadamente no contexto do Estado Social.

Palavras-chave: Projeções demográficas, Modelo de projeção multiestado, População portuguesa, Educação.

Abstract

Access to accurate information on the population structure at a given time, its evolution in recent years and how certain demographic variables may change in the future plays a key role in decision making in several sectors. Educational attainment is an attribute that leads to a greater distinction between members of a population; in this context, the main objective of this work was to produce demographic projections for the Portuguese population by sex, five-year age groups and education level, for the years 2001-2031. To this purpose, we used the multistate projection model.

A multistate model is a model for a stochastic process which, at any time, occupies one of a set of discrete states defined by the group of characteristics considered. Besides the estimation of transition probabilities between levels of education, by sex and age group, it was also needed to consider the existence of differentials regarding fertility, mortality and migration, by level of education. The attainment of effective population at the end of each five-year projection period was based on the multistate cohort component method, embodied in the use of a block Leslie matrix. We considered two scenarios of evolution, one in which educational attainment observed between 2001 and 2006 remains constant in the remainder of the projection period and another in which educational attainment will follow, in the coming years, the trend observed over the last decade. In this second scenario, we modeled the trend in the state proportion for each sex and age group using continuation ratio models.

The results show an increase in the proportion of individuals who complete higher educational levels in all age groups of both sexes. According to the trend scenario, the percentage of individuals aged 65 years or more with a full course of higher education, which stood at 2.4% in 2001, is expected to increase to 11.3% in 2031. Among women, only 7.3% had completed some level of higher education in 2001, a figure that will rise to approximately 22% in 2031, whereas among men this value was only 5.7% in the baseline year and will also rise by 2031, reaching 13.2%. Given that the transition to a higher level of education is an irreversible event, we can expect the proportion of elderly people with higher education levels to continue to rise as the education of younger cohorts seems to evolve positively.

It is intended that this work will serve as a basis to study how the aging population and the rising levels of education can contribute to the planning and monitoring of public policies, particularly in the context of the welfare state.

Keywords: Demographic projections, Multistate projection model, Portuguese population, Education.

Agradecimentos

Dirijo os meus sinceros agradecimentos a todos os que contribuíram para a elaboração deste trabalho:

- À minha orientadora, Prof.^a Doutora Maria Isabel Barão, pela disponibilidade e solicitude, combinadas em conversas da maior relevância em momentos chave;
- À minha coorientadora, Prof.^a Doutora Maria do Rosário Martins, pelo estímulo transmitido desde os tempos de licenciatura, pelas diretrizes sobre a definição do tema e objetivos do trabalho e por todo o apoio e motivação;
- À Prof.^a Doutora Teresa Rodrigues, pela possibilidade de integrar a sua Equipa de Investigação, pelas valiosas sugestões e material disponibilizado;
- Ao Prof. Doutor Paulo Ferrinho, pela oportunidade de realizar o meu trabalho no IHMT-UNL;
- Aos colegas de mestrado, pela troca de conhecimentos e pelo suporte mútuo ao longo dos últimos dois anos; agradeço em especial à Mónica, com quem aprendi o que é ser uma colega extraordinária;
- Aos meus pais, Carolino e Mila, que me ensinaram que o trabalho e dedicação constituem o verdadeiro caminho na prossecução de qualquer objetivo, pelo apoio e incentivo permanentes; à minha restante família, porque me dão a confiança de ter um porto de retorno seguro;
- À minha amiga Sofia, pelo carinho e cumplicidade que nutrem cada encontro, e à minha querida Vera, um agradecimento pelas conversas que despertam a vontade de questionar, investigar, pensar, e um pedido de desculpa pelas ausências repetidas;
- Ao Pedro, pela generosidade de cada gesto, pelo humor que ameniza as contrariedades, pelas conversas que dão sentido aos dias e por tudo o que as palavras não conseguem expressar, e à sua família, porque sempre me fizeram sentir em casa.

Índice

	Página
Introdução	1
Capítulo 1. Modelos populacionais	5
1.1. Conceitos básicos	5
1.2. Tabela de mortalidade	7
1.3. Modelos de população estável	10
1.4. Modelos multiestado	11
Capítulo 2. Projeções demográficas	19
2.1. Principais projeções demográficas produzidas a nível nacional e internacional	19
2.2. Técnicas de projeção demográfica	20
2.3. Projeções demográficas por nível de escolaridade	24
Capítulo 3. Caracterização demográfica recente da população portuguesa	31
3.1. Componentes demográficas	31
3.2. Fatores de crescimento e grupos funcionais	37
Capítulo 4. A Educação em Portugal	41
4.1. A estrutura atual do Sistema Educativo Português e reformas em curso	41
4.2. Indicadores de escolarização da população portuguesa	43
4.3. Objetivos e metas para a Educação	47
Capítulo 5. Metodologia de projeção	49
5.1. População de partida	50
5.2. Desenho do modelo de projeção	51
5.3. Tratamento da incerteza: construção de cenários de projeção	63
5.4. Software utilizado na elaboração das projeções	74
Capítulo 6. Resultados	75
6.1. População residente por sexo, grupo etário e nível de escolaridade, 2001-2031	75
Capítulo 7. Discussão	81
7.1. Avaliação de resultados	81
7.2. Limitações	85
7.3. Relevância do trabalho e desenvolvimentos futuros	86
Conclusões	89
Bibliografia	91
Apêndice A. Tabelas auxiliares	101
Apêndice B. Gráficos auxiliares	111
Apêndice C. Expressão matricial para projeção da população feminina para o quinquénio 2001-2006	113

Lista de Figuras

	Página
1.1.1 Diagrama de Lexis	7
2.2.1 Esquema da projeção demográfica por idade e sexo	23
3.1.1 Esperança de vida à nascença, por sexo, 1980-2010	32
3.1.2 Esperança de vida aos 65 anos, por sexo, 1980-2010	32
3.1.3 Taxa de mortalidade infantil, 1970-2010	33
3.1.4 Taxa de mortalidade infantil por nível de escolaridade da mãe, 1996-2010	34
3.1.5 Índice Sintético de Fecundidade e idade média da mãe ao nascimento do primeiro filho, 1972-2010	35
3.1.6 Taxa de Fecundidade, por grupo etário e nível de escolaridade da mãe, 2010	36
3.1.7 Índice Sintético de Fecundidade, por nível de escolaridade da mãe, 1998-2010	36
3.1.8 Saldo Natural e Saldo Migratório, 1992-2010	37
3.1.9 Emigrantes, por nível de instrução, 2000-2003	38
3.1.10 População estrangeira que solicitou estatuto de residente, por nível de instrução, 2000-2006	38
3.2.1 Taxas de Crescimento Natural, Migratório e Efetivo, 1992-2010	39
3.2.2 Grupos Funcionais, 1992-2010	39
3.2.3 Índices de envelhecimento, longevidade e renovação da população em idade ativa, 1971-2010	40
3.2.4 Índices de dependência total, de jovens e de idosos, 1971-2010	40
4.1.1 Organização do Sistema Educativo Português	43
4.2.1 Taxa real de escolarização, por nível de ensino, 1961-2010	44
4.2.2 Aprendizagem ao longo da vida, 1998-2009	44
4.2.3 Número de inscritos e diplomados no Ensino Superior, 2000/2001 a 2009/2010	45
4.2.4 População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 1981	46
4.2.5 População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 1991	46
4.2.6 População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 2001	46
6.1.1 Grupos funcionais, 2001-2031, por cenário	75
6.1.2 Índices-resumo, 2006-2031, Cenário Tendência	76
6.1.3 Número de residentes por nível de escolaridade, 2001-2031, por cenário	76
6.1.4 Número de residentes com 15-64 anos, por nível de escolaridade, 2001-2031, por cenário	77
6.1.5 Número de residentes com 65 ou mais anos, por nível de escolaridade, 2001-2031, Cenário Tendência	77
6.1.6 População portuguesa por grupo etário, sexo e nível de escolaridade, 1 de janeiro de 2001	78

6.1.7	População portuguesa por grupo etário, sexo e nível de escolaridade, 1 de janeiro de 2031, por cenário	78
B.0.1	População residente em Portugal, por sexo, grupo etário e nível de escolaridade, 2001-2031, Cenário Constante	111
B.0.2	População residente em Portugal, por sexo, grupo etário e nível de escolaridade, 2001-2031, Cenário Tendência	112

Lista de Tabelas

	Página
1.4.1 Classificação de eventos	13
3.1.1 Taxa Bruta de Mortalidade, por sexo, 1970-2010	33
3.1.2 Taxa Bruta de Natalidade, 1970-2010	34
5.2.1 Correspondência entre categorias de escolaridade e níveis ISCED	52
5.2.2 Matriz de eventos para o modelo de projeção por escolaridade	52
5.2.3 Funções de mortalidade estimadas para cada sexo e nível de escolaridade	54
5.2.4 Expressões de cálculo da proporção de sobrevivência, para cada sexo e nível de escolaridade, por grupo etário	55
5.3.1 Índice Sintético de Fecundidade, por nível de escolaridade, 2001-2031	65
5.3.2 Taxa de Fecundidade (‰) por grupo etário e nível de escolaridade da mãe, 2001-2031	66
5.3.3 Proporções de sobrevivência por sexo, grupo etário e nível de escolaridade, 2001-2031	68
5.3.4 Saldo migratório por sexo, grupo etário e nível de escolaridade, 2001-2031	70
5.3.5 Proporções de transição entre níveis de escolaridade, por sexo, 2001-2031	73
7.1.1 População residente (N. ^o), por sexo e grupo etário, 2011, comparação de resultados (Censos 2011, Cenário central INE e Cenário tendência)	82
7.1.2 População residente (N. ^o), por sexo e nível de escolaridade mais elevado completo, 2011, comparação de resultados (Censos 2011 e Cenário tendência)	82
7.1.3 População residente (N. ^o), por sexo, grupo etário e nível de escolaridade, 2011, comparação de resultados (Inquérito ao Emprego e Cenário tendência)	83
A.0.1 População residente (N. ^o) por sexo e grupo etário, 31/12/2000	101
A.0.2 Distribuição dos residentes de cada sexo e grupo etário por nível de escolaridade, Censos 2001	101
A.0.3 População de partida por sexo, grupo etário e nível de escolaridade	101
A.0.4 Taxa de mortalidade infantil (‰) por nível de escolaridade da mãe, 2000-2007	101
A.0.5 Quociente de mortalidade entre as idades exatas x e $(x + n) - {}_nq_x$ - por sexo e grupo etário, 2000-2002 a 2005-2007	102
A.0.6 População feminina com 15-49 anos, por grupo etário e nível de escolaridade, 1998-2010	103
A.0.7 Nados-vivos (N. ^o) por grupo etário da mãe e nível de escolaridade da mãe, 1998-2010	104
A.0.8 Relação de masculinidade (N. ^o), grupo etário 0-4 anos, 2001-2006	105
A.0.9 Saldo migratório anual, 2001-2006	105
A.0.10 Distribuição percentual dos migrantes internacionais por sexo, 2000, 2010	106
A.0.11 Distribuição percentual dos migrantes internacionais por grupo etário e sexo, 2000, 2010	107

A.0.12 Taxas de Fecundidade específicas por idade (‰), variante média, 2005-2030	108
A.0.13 Taxas de mortalidade infantil (‰), por sexo, variante média, 2005-2030	108
A.0.14 Esperança de vida à nascença (anos), por sexo, variante média, 2005-2030	108
A.0.15 Saldo migratório anual, 2006-2031	109
A.0.16 População residente em Portugal, por sexo, grupo etário e nível de escolaridade (2001, 2016, 2031), por cenário	110

Introdução

O conhecimento relativo à situação e dinâmica demográfica de um dado país ou região constitui uma mais-valia determinante em todo o processo de planeamento e monitorização de políticas, seja ao nível estratégico, económico, político, social ou ecológico. O acesso a informação de qualidade relativa à estrutura populacional num dado momento, à sua evolução nos anos recentes e à forma como determinadas variáveis ou medidas demográficas poderão variar no futuro desempenha um papel fundamental na tomada de decisão em sectores tão relevantes como a Saúde, a Economia, o Estado Social ou a Educação.

A elaboração de projecções demográficas surge com o objectivo de obter estimativas da dimensão total ou da composição de uma população no futuro e representa, assim, uma ferramenta essencial em diversos campos (O'Neill et al., 2001; OCDE, 2001). Recentemente, o número e a qualidade das análises prospectivas da população têm aumentado de forma considerável, graças, em parte, ao progresso tecnológico verificado nas últimas décadas, particularmente em termos computacionais (Wilson e Rees, 2005). Tais avanços têm permitido, por exemplo, aperfeiçoar técnicas que permitem considerar a variabilidade presente entre os indivíduos que compõem a população em estudo, o que conduz, regra geral, a resultados mais precisos e completos, comparativamente ao caso em que apenas se estuda o conjunto da população como um todo (O'Neill et al., 2001). Como tal, é necessário determinar que características, de entre as que são relevantes para distinguir os indivíduos, devem ser explicitamente consideradas nas projecções populacionais. A idade e o sexo são os atributos mais frequentemente considerados nas principais projecções produzidas a nível nacional e internacional (ONU, 2011b; INE, 2009). Porém, desde que esteja garantida a existência de dados de boa qualidade, é possível efectuar projecções mais elaboradas, em que se consideram outros atributos, tais como o estado civil, nível de escolaridade, situação profissional, classificação do local de residência (urbano/rural), tipo de agregado familiar ou nível de rendimentos (Willekens, 2006). Estudar cada uma destas dimensões pode ter interesse próprio, caso se pretenda responder a questões específicas (tem-se verificado, por exemplo, que o aumento de agregados familiares unipessoais se associa a um crescente consumo de recursos energéticos e ambientais (Alho e Keilman, 2010) e, portanto, pode ser útil, para áreas como a Economia ou o Ambiente, estudar a estrutura populacional futura por tipo de agregado familiar) permitindo, para além disso, aperfeiçoar os resultados da projecção da dimensão populacional total, na medida em que tais características se associam, muitas vezes, a diferentes comportamentos relativos à mortalidade, fecundidade e migrações (Lutz e KC, 2010). O seu envolvimento na projecção deverá, pois, contribuir para minimizar o eventual viés originado pela heterogeneidade presente entre os indivíduos.

Os métodos demográficos multiestado permitem descrever o desenvolvimento de uma população ao longo do tempo e constituem, actualmente, a metodologia padrão nos casos em que se pretenda considerar uma população dividida de acordo com várias características e as interações entre estes subgrupos (Willekens, 2006). Sendo indicado para o estudo de uma população sujeita a uma divisão multidimensional, em que se consideram vários atributos dos seus elementos, pode, por esse motivo, designar-se também como modelo de projecção multidimensional (Imhoff e Keilman, 1991). Relativamente a cada indivíduo, alguns atributos mantêm-se fixos, como é o caso do sexo, mas a maioria sofre alterações ao longo do tempo, conduzindo a alterações na composição da população. Quando determinado atributo se modifica, ocorre um evento. Ao focar-se sobre os eventos ocorridos no decorrer de um certo período, para um determinado grupo de indivíduos, o modelo de projecção multiestado incorpora uma perspectiva dinâmica da evolução demográfica. Um ponto essencial é que, se a alteração da idade de um indivíduo é inerente à passagem do tempo (desde que o indivíduo se mantenha vivo), a alteração de outras variáveis, como o nível de escolaridade, pelo contrário, pode ocorrer segundo diferentes padrões para indivíduos distintos, e daí a necessidade de estudar em que momento e com que intensidade ocorrem estes eventos, de modo a conseguir estimar a estrutura populacional futura. É evidente que, ao considerar uma combinação de várias características, possibilita-se um estudo mais aprofundado do seu efeito sobre o comportamento demográfico mas, simultaneamente, aumenta-se de forma significativa a complexidade dos dados e das metodologias necessárias.

O nível de instrução é, de facto, um dos atributos que levam a maior distinção entre os membros de uma população. Extensamente reconhecida como um impulsor essencial do desenvolvimento, tanto ao nível individual como da sociedade, a escolaridade exerce influência sobre os comportamentos relativos à fecundidade, mortalidade e migrações (Lutz e KC, 2010; KC et al., 2010; Goujon, 2008). Adicionalmente, associa-se a outras variáveis: por exemplo, em Portugal, os dados fornecidos pelo Inquérito Nacional de Saúde (2005/2006) apontam para que os indivíduos com níveis de escolaridade mais elevados tendam a avaliar o seu estado de saúde de forma mais positiva, em comparação com os menos escolarizados (Henriques e Rodrigues, 2010). A percepção crescente da importância da escolaridade e do capital humano fomentou, nas últimas décadas, a aplicação de métodos de projecção para estudar a composição populacional por nível de escolaridade, em vários países (Ahuja e Filmer, 1995b; Goujon e Wils, 1996; Lutz, 1999; Lutz e Goujon, 2001; Lutz et al., 2005; Goujon et al., 2007; KC et al., 2010).

Neste contexto, o objectivo central do presente trabalho consistiu na elaboração de projecções demográficas para a população portuguesa por sexo, grupo etário e nível de escolaridade, para o período 2001-2031. Para tal, recorreu-se ao modelo de projecção multiestado e à definição de diferentes cenários de evolução, relativos tanto às componentes demográficas fecundidade, mortalidade e migrações, como aos níveis de escolarização. Os dados censitários de 2011 disponíveis até ao momento, os resultados das projecções demográficas elaboradas pelo Instituto Nacional de Estatística (INE), I.P., e as estimativas obtidas no Inquérito ao Emprego (INE) serviram de referência para a avaliação dos

resultados do exercício de projecção. Com vista à concretização deste objectivo principal, foram definidos os seguintes objectivos específicos: a) caracterizar a evolução demográfica recente da população portuguesa; b) caracterizar genericamente os principais tipos de projecções demográficas existentes e analisar a importância de projectar a população por nível de escolaridade; c) descrever o modelo de projecção multiestado; d) estimar os parâmetros de interesse e estabelecer possíveis cenários de evolução e, por fim, f) analisar e discutir os resultados obtidos. Este trabalho surgiu no âmbito do projecto "Envelhecimento e Saúde em Portugal: Políticas e Práticas", PTDC/CS-DEM/109967/2009, co-financiado pela Fundação para a Ciência e Tecnologia e pelo Programa Operacional Temático Factores de Competitividade - COMPETE, em que se pretende avaliar o impacto do envelhecimento populacional sobre as políticas e práticas de saúde em Portugal numa óptica prospectiva, associado ao estudo da relação entre o estado de saúde e o aumento esperado dos níveis de instrução da população portuguesa.

Esta tese encontra-se organizada em sete capítulos principais, que incluem uma breve descrição sobre os principais modelos populacionais (capítulo 1), nos quais se baseiam as técnicas de projecção demográfica comumente utilizadas (capítulo 2), apresentando-se depois a caracterização demográfica recente da população portuguesa (capítulo 3) e a descrição das características essenciais do Sistema Educativo Português (capítulo 4), cujo conhecimento é essencial na definição da metodologia de projecção (capítulo 5). Os dois últimos capítulos centram-se na apresentação dos resultados obtidos (capítulo 6) e na discussão dos mesmos, bem como das principais limitações surgidas no decorrer do trabalho efectuado (capítulo 7). Por último, apontam-se as conclusões e considerações finais.

Modelos populacionais

Os modelos populacionais pretendem expressar de que forma os processos demográficos centrais – fecundidade, mortalidade e movimentos migratórios – determinam a dimensão e estrutura de uma dada população, revelando as dinâmicas subjacentes de crescimento e recessão, concentração e dispersão, envelhecimento e renovação (Schoen, 2006d). Desde a primeira tabela de mortalidade produzida por John Graunt, no século XVII, a complexidade dos modelos demográficos tem aumentado notavelmente, permitindo capturar e expressar o comportamento populacional com um rigor crescente. Neste capítulo, pretende dar-se uma visão inicial integrada sobre os principais conceitos utilizados em Demografia, prosseguindo-se com a descrição dos principais modelos populacionais, os modelos de população estável, e a sua extensão para os modelos multiestado.

1.1. Conceitos básicos

A alteração da dimensão de uma população ao longo de determinado período resulta do balanço entre as entradas e as saídas de indivíduos verificadas no mesmo intervalo temporal. De um modo geral, considerando a população residente em determinado país ou região, é através dos nascimentos ou das imigrações que se processam as entradas na população; por outro lado, as mortes e emigrações constituem as vias de saída¹. Este processo pode ser esquematizado através da **Equação de Equilíbrio da Variação Populacional**, dada pela seguinte expressão:

$$P[t] = P[0] + N[0, t] - O[0, t] + I[0, t] - E[0, t]$$

onde: t é um instante de tempo, geralmente medido em anos; $P[t]$ e $P[0]$ correspondem, respetivamente, ao número de elementos da população no início e no final do intervalo; $N[0, t]$ é o número de nascimentos, $O[0, t]$ o número de óbitos, $I[0, t]$ o número de imigrantes e $E[0, t]$ o número de emigrantes no intervalo $[0, t]$. Esta equação pode servir de base à estimação da dimensão populacional numa data futura, através da definição de pressupostos sobre a ocorrência de nascimentos, mortes e migrações até esse instante. Muitas vezes, e sobretudo no caso da fecundidade e da mortalidade, estas hipóteses

¹Consoante a definição específica da população em estudo, é possível que alguma destas formas de entrada ou saída não esteja presente.

são estabelecidas em termos de **taxas** e não sob a forma do número absoluto de nascimentos ou mortes², sendo que a natureza e estimação destas taxas constitui uma parte importante da metodologia demográfica (Shryock et al., 1973). As taxas de mortalidade ou de fecundidade referem-se, regra geral, ao número esperado de ocorrências (mortes ou nascimentos), por pessoa e por ano de exposição ao risco, no período de tempo considerado. São, portanto, designadas por taxas de ocorrência/exposição (Schoen, 2006d; Carvalho, 2010). Se nos referirmos à ocorrência de nascimentos ou mortes num pequeno intervalo de tempo, digamos $[t, t + \Delta t]$ com $\Delta t \rightarrow 0$, trabalhamos com uma medida do risco instantâneo de ocorrer um nascimento ou uma morte no instante exato t . Essa medida, mais precisa que a taxa correspondente, mas não diretamente observável, pode designar-se por força, intensidade ou **taxa instantânea** de nascimento ou de morte (Schoen, 2006d).

Na análise demográfica, os indivíduos e os eventos têm de ser posicionados numa determinada escala de tempo, sendo a idade (tempo individual) e o tempo de calendário (tempo histórico) as escalas habitualmente utilizadas. Como tal, podem distinguir-se duas perspetivas principais para abordar a ocorrência de eventos demográficos: a de coorte e a de período. Numa determinada **coorte** estão agrupados os indivíduos que experimentaram um mesmo acontecimento demográfico no mesmo intervalo de tempo (Ryder, 1965; Willekens, 2006). Se o nascimento é o acontecimento que define a coorte, o grupo de indivíduos que nasceu durante o mesmo período constitui uma **coorte de nascimento**. É comum subdividir a população considerando coortes de nascimento, já que, à partida, os indivíduos nascidos durante o mesmo período partilham características que dão a cada coorte um carácter distinto. Nesse sentido, apesar de os membros de uma mesma coorte terem histórias de vida distintas, os parâmetros que definem a trajetória de vida são provavelmente parecidos, quando comparados com os de membros de uma coorte diferente. Podemos então registar os eventos ocorridos pelos elementos de uma mesma coorte ao longo do tempo e calcular as respetivas taxas específicas por coorte. Por outro lado, referimo-nos a um **período** como um intervalo de tempo de calendário relativamente pequeno: o número de mortes ocorridas durante um dado ano constitui um exemplo de dado de período. Num qualquer período, um certo número de coortes compõe a população e cada uma delas contribui para os eventos verificados nesse intervalo. É esta importante interacção entre as duas perspetivas que serve de base ao **esquema período-coorte**, perspetiva central na elaboração de projecções populacionais: a dimensão e estrutura de uma população são estimadas para diferentes instantes t no tempo de calendário, abrangendo os indivíduos pertencentes a vários subgrupos, geralmente definidos pelo sexo e grupo etário. Deste modo, o tempo de calendário é habitualmente tratado como uma variável contínua e a idade como uma variável discreta, visto ser considerada sob a forma de grupos etários, geralmente anuais ou quinquenais (Willekens, 2006). A estimação de taxas demográficas específicas por período e coorte conduz à possibilidade de aplicar a Equação de Equilíbrio da Variação Populacional a cada subgrupo da população, para cada um dos períodos considerados. O **Diagrama de Lexis** permite

²A descrição das migrações externas, por outro lado, baseia-se habitualmente na expressão conjunta das emigrações e imigrações em termos de números absolutos, dada a dificuldade associada à definição da população em risco de imigração (Alho e Keilman, 2010).

esquematizar a ocorrência de eventos demográficos (em particular, da mortalidade) ao longo do tempo, enquadrados de acordo com as duas escalas, tempo de calendário (eixo dos XX) e idade (eixo dos YY), usando a mesma unidade de tempo em ambos os eixos. Cada linha a 45° representa a vida de um indivíduo e o avanço de cada coorte ao longo do tempo é delimitado pelas duas linhas com origem no início e no final do intervalo de tempo a que a coorte diz respeito. Os tempos de vida dos elementos de uma coorte podem ser registados neste diagrama, permitindo distinguir os segmentos que devem ser utilizados na construção do tempo de exposição para cada coorte num dado período. No entanto, se as linhas de vida individuais exatas não forem conhecidas, são necessárias estimativas do tempo de exposição em pessoa-anos, baseadas no número de elementos de cada coorte no início e no final de cada intervalo de tempo, de modo a poderem ser calculadas as taxas necessárias (Willekens, 2006; Carvalho, 2010).

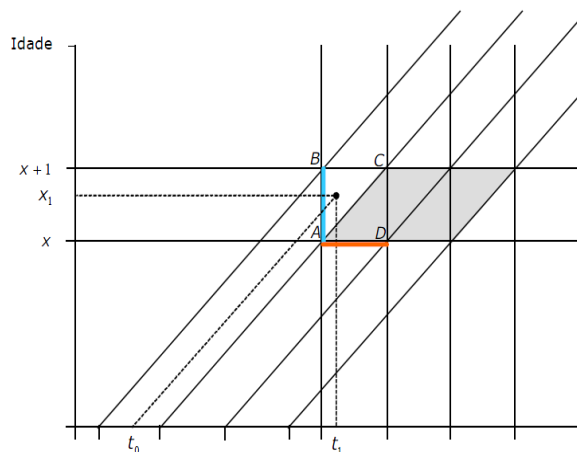


FIGURA 1.1.1. Diagrama de Lexis

É representada a linha de vida de um indivíduo que nasce no momento t_0 e morre no momento t_1 à idade x_1 , entre as idades exatas x e $x + 1$. O segmento de reta vertical a azul representa o número de indivíduos que no momento t têm x anos de idade; O segmento de reta horizontal a vermelho representa o número de indivíduos que no ano t cumpriram o x -ésimo aniversário. No decurso do ano t , os óbitos entre os indivíduos com x anos inteiros vividos, estão localizados num quadrado delimitado pelas coordenadas ABCD.

Fonte: INE (2010)

1.2. Tabela de mortalidade

A experiência de mortalidade de uma coorte de nascimento pode ser sumarizada com base numa tabela de mortalidade que, na sua forma mais clássica, traduz toda a informação relativa ao desaparecimento por morte do conjunto de elementos de uma coorte, ao longo do tempo (Carvalho, 2010). Na prática, é geralmente o cálculo das taxas de mortalidade específicas por idade para um determinado período que serve de base à construção de uma tabela de mortalidade que, não sendo referente a uma coorte real, reflete o fenómeno de mortalidade de uma coorte hipotética, sujeita às condições de mortalidade observadas no período em questão. As funções desta tabela de mortalidade corrente, também designada por contemporânea, referem-se a uma população estacionária, em que a distribuição dos elementos à medida que a idade avança resulta das taxas de mortalidade calculadas sobre a coorte hipotética, sem que ocorram entradas ou saídas por migração externa. Esta coorte tem origem, regra

geral, em 100.000 novos nascimentos, valor habitualmente designado por raiz ou radical da tabela e que se mantém constante. Numa tabela de mortalidade são geralmente apresentados os seguintes valores (INE, 2010; Carvalho, 2010):

ℓ_x :	Número de sobreviventes à idade exata x
${}_n d_x$:	Número de mortes entre as idades x e $x + n$
${}_n m_x$:	Taxa de mortalidade da coorte entre as idades x e $x + n$
${}_n a_x$:	Número médio de pessoa-anos vividos no intervalo $[x, x+n[$ pelos indivíduos que morreram com idade nesse intervalo
${}_n p_x$:	Probabilidade de um indivíduo que atingiu a idade x sobreviver até à idade $x + n$
${}_n q_x$:	Probabilidade de um indivíduo que atingiu a idade x falecer antes de atingir a idade $x + n$
${}_n L_x$:	Pessoa-anos vividos entre as idades x e $x + n$
T_x :	Pessoa-anos vividos acima da idade x
e_x :	Esperança de vida completa à idade x

Deve notar-se que ${}_n d_x$, ${}_n m_x$, ${}_n a_x$, ${}_n p_x$, ${}_n q_x$ e ${}_n L_x$ se referem a intervalos de idade³ e ℓ_x , T_x e e_x são relativos a idades exatas.

A principal dificuldade associada ao cálculo de uma tabela de mortalidade corrente consiste na transformação das taxas de mortalidade por idade observadas no período em questão, designadas por ${}_n M_x$, nas respetivas probabilidades de morte por idade, ${}_n q_x$. Numa coorte verdadeira, a taxa de mortalidade e a probabilidade de morte específicas são dadas, respetivamente, por:

$${}_n m_x = \frac{{}_n d_x}{{}_n L_x}$$

e

$${}_n q_x = \frac{{}_n d_x}{\ell_x}$$

Por outro lado:

$${}_n L_x = n \cdot \ell_{x+n} + {}_n A_x$$

onde ${}_n A_x$ corresponde ao número de pessoa-anos vividos pelos elementos da coorte que morreram no intervalo $[x, x+n[$. Dado que ${}_n a_x$ representa o número médio de pessoa-anos vividos pelos elementos da coorte que morreram nesse intervalo, com ${}_n a_x = \frac{{}_n A_x}{{}_n d_x}$, vem que (Carvalho, 2010):

$$\begin{aligned} {}_n L_x &= n \cdot \ell_{x+n} + {}_n a_x \cdot {}_n d_x = n(\ell_x - {}_n d_x) + {}_n a_x \cdot {}_n d_x \Leftrightarrow \\ &\Leftrightarrow n \cdot \ell_x = {}_n L_x + (n - {}_n a_x) {}_n d_x \Leftrightarrow \\ &\Leftrightarrow \ell_x = \frac{1}{n} [{}_n L_x + (n - {}_n a_x) {}_n d_x] \end{aligned}$$

³A amplitude dos intervalos etários considerados na tabela de mortalidade pode variar: quando se considera a idade ano a ano referimo-nos a tabelas de mortalidade completas; numa tabela abreviada, é habitual o primeiro intervalo ter amplitude de 1 ano, o segundo 4 anos e os restantes 5 anos (INE, 2010; Carvalho, 2010).

Como tal:

$$\begin{aligned} {}_nq_x &= \frac{{}_nd_x}{\ell_x} = \frac{{}_nd_x}{\ell_x} = \frac{n \cdot {}_nd_x}{{}_nL_x + (n - {}_na_x){}_nd_x} = \frac{n \cdot {}_nd_x / {}_nL_x}{1 + (n - {}_na_x){}_nd_x / {}_nL_x} \Leftrightarrow \\ &\Leftrightarrow {}_nq_x = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x){}_nm_x} \end{aligned}$$

verificando-se que a transformação de ${}_nm_x$ em ${}_nq_x$ depende exclusivamente da função ${}_na_x$. Assim, visto que na tabela de mortalidade corrente se admite que a coorte hipotética experimenta a mortalidade específica relativa à coorte verdadeira, é somente necessário obter um conjunto de ${}_na_x$ para proceder à sua construção. Para tal, existem diversas abordagens possíveis como, por exemplo: a) aproximar ℓ_x entre x e $x + n$ por uma função linear, o que equivale a assumir que as mortes se distribuem uniformemente ao longo do intervalo $[x, x + n[$ ou ocorrem exatamente a meio do intervalo e, como tal, ${}_na_x = \frac{n}{2}$; b) utilizar a aproximação exponencial de ℓ_x entre x e $x + n$; c) observar ${}_na_x$ diretamente na população em estudo, quando se conhecem as idades exatas de morte dos seus elementos; ou d) substituir os verdadeiros valores de ${}_na_x$ por valores calculados para uma coorte de uma outra população (Carvalho, 2010). Obtidos os valores de ${}_na_x$ pelo método selecionado, é possível concluir a construção da tabela de mortalidade corrente.

Levando em consideração que, apesar de a tabela de mortalidade fornecer funções da mortalidade relativas a idades inteiras ou a intervalos com extremos inteiros, o risco de mortalidade atua continuamente sobre uma coorte, é importante obter uma função de sobrevivência, $S(x)$, que expressa a probabilidade de sobrevivência dos elementos da coorte à medida que o tempo decorre, num ponto de vista contínuo (Shryock et al., 1973; Carvalho, 2010). Em termos formais:

$$S(x) = P[I \geq x]$$

onde I corresponde ao tempo de vida, definido como o intervalo desde que um indivíduo nasce até à sua morte.

Se $S(x)$ for conhecida, esta função pode ser utilizada para construir uma tabela de mortalidade, partindo de um dado radical, ℓ_0 , tabelando, para valores inteiros de x , a função $\ell(x) = \ell_0 S(x)$, que representa o número médio de sobreviventes à idade exata x ⁴. No entanto, $S(x)$ é normalmente desconhecida, encontrando-se disponíveis, pelo contrário, os valores de $\ell(x) = \ell_x$. Desta forma, pode recorrer-se à tabela de mortalidade para obter estimativas de $S(x)$, que podemos designar como probabilidades de sobrevivência, sendo:

$$S(x) = \frac{\ell_x}{\ell_0} = {}_xp_0 \quad (1)$$

As probabilidades de sobrevivência podem ser específicas para determinadas características que determinam subgrupos da população.

⁴Em termos probabilísticos, $\ell(x)$ é uma variável aleatória binomial de valor médio $\ell_0 S(x)$

1.3. Modelos de população estável

Os modelos baseados no conceito de população estável foram introduzidos por Sharpe e Lotka (1911) e permaneceram dominantes na matemática demográfica do século XX. Numa população estável, as taxas de fertilidade e de mortalidade por idade são constantes e a população é fechada para migrações. Em consequência, origina-se, num prazo suficientemente longo, uma população que cresce exponencialmente ao longo do tempo, e cuja composição etária se mantém inalterada. A população estacionária surge como um caso particular da população estável, em que o fluxo de nascimentos é constante e, como tal, a taxa de crescimento intrínseca é nula, com a manutenção dos efetivos totais e por classe etária. As populações estáveis diferem das populações estacionárias num ponto fundamental: numa população estável, as medidas de período não são as mesmas que as medidas de coorte. O comportamento das várias coortes de uma população estável é o mesmo (as taxas mantêm-se constantes), mas a sua dimensão é distinta, dado que o número de nascimentos por período não se mantém. Por outro lado, mantém-se a estrutura da população em cada período, mas não a sua dimensão, já que em cada período a população reflete a combinação de coortes de diferentes dimensões (Schoen, 2006d).

As representações discretas de populações estáveis foram propostas inicialmente por Bernardelli (1941), posteriormente por Lewis (1942) e, mais tarde, por Leslie (1945), passando o modelo matricial resultante a ser conhecido por Modelo de Leslie. Com base na população feminina⁵, dividida em $m + 1$ grupos etários de amplitude n , observada em pontos discretos no tempo, podemos definir a seguinte matriz (Carvalho, 2010):

$$\mathbf{M} = \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{m-1} & b_m \\ s_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & s_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s_{m-1} & 0 \end{bmatrix}$$

Cada b_i traduz a contribuição das pessoas no i -ésimo grupo etário no momento inicial t para o número de pessoas no primeiro grupo etário no final do intervalo, em $t + n$ (correspondendo, neste caso, ao número médio de filhas nascidas por mulher do grupo etário x no intervalo $[t, t + n[$ e que sobrevivem até $t + n$). Os rácios de sobrevivência na subdiagonal, s_i , refletem a proporção de pessoas no i -ésimo grupo etário no momento inicial t que sobrevivem de modo a pertencer ao grupo etário $i + 1$ no momento $t + n$. Estas taxas são baseadas no número de sobreviventes entre cada intervalo de idades por cada ano de risco de sobrevivência, sendo calculadas através da expressão:

$$s_i = \frac{L_{i+1}}{L_i} \quad (2)$$

⁵O estudo da fecundidade é baseado habitualmente na população feminina, pois é mais fácil associar os nascimentos às mães e o intervalo de idade reprodutiva nas mulheres é mais curto e definido com maior clareza (habitualmente, considera-se que a idade reprodutiva de uma mulher se situa entre os 15 e os 49 anos).

onde L_i é o número de pessoa-anos vividos entre as idades que delimitam o grupo etário i . Notemos a diferença existente entre as probabilidades de sobrevivência $S(x)$, cuja expressão de cálculo foi apresentada na secção anterior (equação 1), e as taxas de sobrevivência s_i : as primeiras são relativas a idades exatas, enquanto que as últimas se referem a grupos etários; por outro lado, cada s_i reflete a sobrevivência condicional para o grupo etário e momento seguinte, dada a sobrevivência no instante inicial, e baseia-se em taxas de ocorrência/exposição, L_i , em que é descontado o tempo em que os indivíduos que, por morrerem, deixaram de estar em risco de sobreviver, para calcular o tempo de risco da população; pelo contrário, numa probabilidade de sobrevivência, não condicional, utiliza-se a proporção de indivíduos que sobreviveram entre determinadas idades, digamos 0 e x , relativamente ao total de indivíduos no momento inicial (com idade 0).

Uma população estável pode ser caracterizada por uma matriz de Leslie que seja constante ao longo do tempo, e a estrutura da matriz permite estimar a taxa de crescimento estável e a estrutura etária correspondentes. Se restringirmos a matriz \mathbf{M} aos grupos etários inferiores à idade considerada como limite superior para o período reprodutivo, obtemos uma matriz que podemos designar por \mathbf{A} . A decomposição desta matriz nos seus valores próprios permite conhecer o seu valor próprio principal, habitualmente representado por λ , que corresponde ao parâmetro de crescimento estável da população ao longo do intervalo de amplitude n . Como tal, $\lambda = \exp(nr)$, onde r é a taxa de crescimento intrínseca. O vetor próprio principal à direita de \mathbf{A} , que passa a ser designado por \mathbf{u} , tem o valor 1 como primeiro elemento e u_i na i -ésima posição. Na população estável, u_i descreve a composição etária da população, ao fornecer a dimensão do i -ésimo grupo etário relativamente ao primeiro grupo etário. Algebricamente, para $i > 1$, $u_i = p_i \lambda^{1-i}$, onde $p_i = \prod_{j=1}^{i-1} s_j$ indica a probabilidade de sobreviver do primeiro ao i -ésimo grupo etário.

1.4. Modelos multiestado

Os modelos multiestado apresentam-se como uma extensão dos modelos que descrevem a dinâmica populacional por idade e sexo, permitindo uma análise demográfica que considere, adicionalmente, outras características. Um modelo multiestado pode descrever a trajetória de vida de uma coorte, por exemplo, com recurso a uma tabela de mortalidade multiestado: enquanto a tabela de mortalidade tradicional descreve somente a experiência de mortalidade, a tabela multiestado descreve, para além disso, a experiência de mobilidade de uma coorte (Willekens, 2006).

Um modelo multiestado é definido como um modelo para um processo estocástico, que num dado instante ocupa um de entre um conjunto de estados discretos, definidos pelo conjunto de características consideradas. Uma mudança de estado designa-se por transição ou evento. A estrutura dos estados especifica que estados são considerados e que transições entre estados são possíveis. Em termos estatísticos, o modelo especifica a estrutura dos estados e a forma da função de risco para cada transição possível (Hougaard, 1999).

Uma característica central destes modelos é o facto de se basearem no recurso a matrizes, referentes, na sua versão mais simples, à ocorrência de eventos em tempo discreto, que permitem, partindo de uma população inicial, projetar a estrutura populacional ao longo do tempo através de métodos convencionais de multiplicação matricial. Tais matrizes são designadas por matrizes de projeção populacional (em inglês, *population projection matrix*, ou simplesmente PPM). A sua forma depende do modelo populacional selecionado que, por sua vez, é condicionado pela natureza da população em estudo, pelos objetivos da projeção e pelos dados de partida que se encontrem disponíveis.

1.4.1. Curso de vida como um processo estocástico.

1.4.1.1. *Estados e espaço dos estados.* De uma forma geral, os modelos demográficos multiestado baseiam-se numa expansão multidimensional da tabela de mortalidade ou dos modelos de população estável: a população classificada de acordo com a idade e sexo é dividida de acordo com um determinado número de estados, que correspondem ao conjunto exaustivo e mutuamente exclusivo dos atributos ou combinações de atributos associados a uma ou mais características demográficas adicionais. O conjunto de todos os estados considerados compõe o espaço dos estados, que se assume ser finito, e que passaremos a representar por $E = \{1, 2, 3, \dots, k\}$, sendo k a dimensão do espaço dos estados. Ao vetor no espaço de estados cujos elementos consistem no número de indivíduos em cada estado num determinado ponto no tempo, designamos por vetor de estados; a cada sexo, grupo etário e ponto no tempo corresponde um vetor de estados, de dimensão $(k \times 1)$ representado por:

$$\ell_{z,t}^s = \begin{bmatrix} \ell_{z,1,t}^s & \ell_{z,2,t}^s & \dots & \ell_{z,k,t}^s \end{bmatrix}'$$

onde s corresponde ao sexo ($s = 1$ para o sexo feminino e $s = 2$ para o sexo masculino, por exemplo), z é o índice para o grupo etário ($z = 1, \dots, m$, sendo m o número de grupos etários considerados e $z = 1$ refere-se ao grupo etário dos indivíduos nascidos durante o intervalo de tempo anterior a t) e t é o instante. A população de cada sexo num dado instante pode ser então representada pelo vetor:

$$\ell_t^s = \begin{bmatrix} \ell_{1,t}^s & \ell_{2,t}^s & \ell_{3,t}^s & \dots & \ell_{m,t}^s \end{bmatrix}'$$

1.4.1.2. *Eventos.* Como referido anteriormente, a modificação de um atributo designa-se por evento. Admitindo que a população não é fechada – alguns indivíduos deixam a população, através da morte ou emigração, e outros entram, por via do nascimento ou imigração – podemos distinguir dois tipos de eventos: eventos internos, que correspondem a transições entre estados (no caso concreto do nível de escolaridade, por exemplo, referimo-nos à conclusão de um determinado grau de ensino: o indivíduo passa para o estado referente ao nível de ensino que concluiu) e eventos externos, que compreendem as entradas e saídas da população (Imhoff e Keilman, 1991; Willekens, 2006). As entradas na população podem ainda ser distinguidas consoante a origem: o nascimento é designado por entrada endógena e a imigração por entrada exógena (tabela 5.2.2). Considera-se que a saída do sistema (por

morte ou emigração) para os indivíduos que acabaram de entrar (por nascimento ou imigração) é um evento sem relevância, já que estes indivíduos não chegam a entrar, de facto, no sistema. Os eventos conduzem, desta forma, a alterações na estrutura da população e, em resultado, o desenvolvimento demográfico pode ser descrito em termos dos eventos ocorridos no período em estudo.

TABELA 1.4.1. *Classificação de eventos*

Fonte: Adaptado de Imhoff e Keilman (1991)

			Posição após evento	
			Posições internas	Posições externas
			$1, 2, 3, \dots, k$	Morte/Resto do Mundo
Posição antes do evento	Posições internas	$1, 2, 3, \dots, k$	Eventos internos	Saídas
	Posições externas	Não nascido	Entradas endógenas	Não relevante
		Resto do mundo	Entradas exógenas	Não relevante

1.4.1.3. *A biografia individual.* Cada um dos indivíduos que compõem a população tem um percurso de vida próprio: ao entrar na população, através do nascimento ou da imigração, ocupa um determinado estado e, à medida que envelhece, pode transitar para outros estados, permanecendo em cada um deles durante um certo período de tempo. A família de variáveis aleatórias que indicam o estado ocupado pelo indivíduo em instantes consecutivos desde a entrada até a saída da população é, portanto, um processo estocástico. Assim, esta sequência de variáveis aleatórias, representadas por $\{X(t), t \geq 0\}$, descreve o curso de vida do indivíduo (Willekens, 2006).

Num determinado instante t , a trajetória do processo até t é, em princípio, conhecida (ou seja, sabemos que estados é que o indivíduo ocupou até esse momento). Desconhecemos, contudo, as realizações das variáveis aleatórias relativas a momentos posteriores a t . Assim, temos de recorrer à distribuição de probabilidade dos possíveis valores de cada variável para determinar quais os valores mais plausíveis e, em particular, qual o seu valor esperado, dada a informação que estiver disponível, tal como os eventos ocorridos antes de t , a idade e sexo do indivíduo, ou o momento de calendário (Willekens, 2006). Uma abordagem comum consiste em assumir que, no que respeita à trajetória de vida já observada, somente o estado ocupado no instante t contribui para explicar o comportamento da variável aleatória referente a um instante seguinte (caso em que a probabilidade do processo assumir um qualquer comportamento futuro não é alterada pelo conhecimento sobre os estados ocupados nos instantes anteriores a t); estamos, consequentemente, perante um processo de Markov com espaço de estados finito (Imhoff e Keilman, 1991; Willekens, 2006; Müller, 2007). Neste contexto, podemos definir a função que fornece a probabilidade de um indivíduo que ocupa o estado i no instante t ocupar

o estado j no instante $t + h$ como:

$$p_{ij}(t, t + h) = P[X(t + h) = j | X(t) = i] \quad (3)$$

sendo $h > 0$ e $i, j = 1, \dots, k$. Esta é a chamada função probabilidade de transição (Willekens, 2006; Müller, 2007).

À medida que h tende para zero, os intervalos entre instantes sucessivos ficam cada vez mais estreitos e passamos a medir as transições em tempo contínuo. A uma transição medida num tempo contínuo referimo-nos como uma transição direta entre estados ou salto imediato entre posições. Desta forma, a densidade condicional de um movimento de i para j (com $i \neq j$) dado que o indivíduo está vivo e se encontra no estado i no instante t é:

$$\mu_{ij}(t) = \lim_{h \rightarrow 0} \frac{P[X(t + h) = j | X(t) = i]}{h} = \lim_{h \rightarrow 0} \frac{p_{ij}(t, t + h)}{h} \quad i \neq j \quad (4)$$

O parâmetro $q_{ij}(t)$ é a taxa a que o indivíduo se move de i para j durante o intervalo infinitesimal após o instante t , ou seja, é o limite, quando h se aproxima de zero, da probabilidade de transição de i para j entre t e $t + h$, por unidade de tempo. É conhecida como taxa instantânea de transição ou intensidade de transição no instante t e é o parâmetro de base, que governa o processo estocástico que descreve a trajetória de vida.

Podemos definir também:

$$\mu_{ii}(t) = - \sum_{j \neq i} \mu_{ij}(t) = - \lim_{h \rightarrow 0} \frac{1 - p_{ii}(t, t + h)}{h}$$

Se trabalharmos em tempo contínuo, por admitirmos que os eventos podem ocorrer em qualquer instante, e com um espaço de estados discreto e finito, podemos classificar o curso de vida como uma cadeia de Markov a tempo contínuo (Müller, 2007). As respetivas probabilidades de transição satisfazem a equação de Chapman-Kolmogorov:

$$p_{ij}(t, t + h) = \sum_{r=1}^k p_{ir}(t, t + h - \Delta h) p_{rj}(t + h - \Delta h, t + h)$$

com $h > 0$ e $0 < \Delta h < h$ e para quaisquer estados i e j . Esta equação indica que a probabilidade de transição do estado i para j pode ser decomposta nas passagens por um estado intermédio r . Devido a esta propriedade, podemos escrever, na forma matricial:

$$\mathbf{P}(t, t + h) = \mathbf{P}(t, t + h - \Delta h) \mathbf{P}(t + h - \Delta h, t + h)$$

onde o primeiro termo do lado direito da equação é a matriz de probabilidades de transição em tempo discreto durante o intervalo de t a $t + h - \Delta h$ e o segundo termo é a matriz de probabilidades

de transição durante o intervalo $t + h - \Delta h$ a $t + h$. Se subtrairmos $\mathbf{P}(t, t + h - \Delta h)$ a ambos os lados da equação, ficamos com:

$$\mathbf{P}(t, t + h) - \mathbf{P}(t, t + h - \Delta h) = \mathbf{P}(t, t + h - \Delta h) [\mathbf{P}(t + h - \Delta h, t + h) - \mathbf{I}]$$

Quando $\Delta h \rightarrow 0$, obtemos:

$$\lim_{\Delta h \rightarrow 0} [\mathbf{P}(t, t + h) - \mathbf{P}(t, t + h - \Delta h)] = \mathbf{P}(t, t + h - \Delta h) \lim_{\Delta h \rightarrow 0} [\mathbf{P}(t + h - \Delta h, t + h) - \mathbf{I}]$$

o que corresponde à equação diferencial que descreve a dinâmica do sistema multiestado, que designamos por equação diferencial de Kolmogorov progressiva:

$$\frac{d \mathbf{P}(t, t + h)}{d h} = \mathbf{P}(t, t + h) \mathbf{Q}(t)$$

em que $\mathbf{Q}(t)$ é a matriz das intensidades de transição $\mu_{ij}(t)$.

No contexto da Análise de Sobrevida, $\mu_{ij}(t)$ representa a função de risco (*hazard function*) da transição de i para j , podendo ser igualmente representada por:

$$\mu_{ij}(t) = \lim_{h \rightarrow 0} \frac{P[t \leq T < t + h | T \geq t]}{h} \quad i \neq j$$

onde T corresponde ao instante em que o indivíduo atinge o estado j vindo do estado i (Putter et al., 2007). As intensidades de transição podem ser estimadas a partir de observações sobre o número e momento de ocorrência das transições, sendo que a informação disponível sobre a ocupação de cada estado é muitas vezes incompleta. Assim, podem ser usados métodos de estimação não paramétricos que considerem a existência de censura (que ocorre quando o instante exato de transição não é observado), tal como o método de Kaplan-Meier, o método de Nelson-Aalen ou a construção de uma tabela de sobrevivência. Uma abordagem alternativa é a parametrização da dependência temporal das intensidades de transição, relativamente à idade e/ou ao tempo de calendário. É ainda possível relacionar as intensidades de transição numa determinada idade e ponto no tempo com um conjunto de covariáveis (Willekens, 2006).

Na prática, a utilização de modelos multiestado em demografia e nas ciências atuariais baseia-se geralmente numa das seguintes abordagens: o modelo exponencial ou o modelo linear. O primeiro baseia-se no pressuposto de manutenção, ao longo do tempo, das intensidades de transição entre cada dois estados. O processo homogéneo, em que $\mu_{ij}(t) = \mu_{ij}$, é um processo estacionário: $p_{ij}(t, t + h)$ depende somente de h e podemos escrever $\mathbf{P}(h) = \mathbf{P}(0, h) = \mathbf{P}(t, t + h)$. Neste caso, as equações de Kolmogorov progressivas

$$\frac{d \mathbf{P}(h)}{d h} = \mathbf{P}(h) \mathbf{Q}$$

onde $\mathbf{Q} = (\mu_{ij})$ admitem a única solução:

$$\mathbf{P}(h) = \exp(\mathbf{Q}h) = \sum_{r=0}^{\infty} \frac{\mathbf{Q}^r h^r}{r!}$$

sujeita à condição inicial $\mathbf{P}(0) = \mathbf{I}$ (Cox e Miller, 1965). Neste caso, a sucessão das variáveis aleatórias dos tempos entre eventos é i.i.d. com a distribuição exponencial (Müller, 2007). A escolha deste modelo é frequente na análise estatística da duração até ao evento e nos casos em que se pretende trabalhar em tempo contínuo (Imhoff e Keilman, 1991).

Uma via alternativa passa por estabelecer que todos os eventos são uniformemente distribuídos ao longo do tempo, o que resulta no chamado modelo linear, que é ainda a abordagem mais comum. De acordo com este pressuposto, o número de eventos em cada subintervalo mantém-se constante e é independente do número de sobreviventes e da sua distribuição pelos vários estados no início do subintervalo. Em linguagem estatística, o modelo assume que as funções densidade de probabilidade das diferentes transições⁶ são constantes à medida que o tempo decorre (Imhoff e Keilman, 1991; Willekens, 2006).

1.4.1.4. A biografia de coorte. As diferenças nas biografias individuais entre os elementos de uma mesma coorte podem ser atribuídas a diferenças nos atributos pessoais, experiências e circunstâncias, e ao efeito do acaso. Se assumirmos que o acaso é o único fator que explica tais diferenças, todos os membros da mesma coorte terão o mesmo valor esperado dos parâmetros que governam a trajetória de vida. Por outro lado, assume-se a ausência de interações entre os indivíduos de forma a assumir a independência das trajetórias de vida individuais que se conjugam na trajetória de vida de uma coorte. A biografia de coorte pode ser, tal como a biografia individual, caracterizada por uma cadeia de Markov de tempo contínuo e com os mesmos parâmetros fundamentais: as intensidades de transição entre i e j para cada instante t , representadas agora por $\mu_{ij}(z, t)$ por serem específicas para os indivíduos do grupo etário z , pertencentes a uma mesma coorte de nascimento.

1.4.2. Tabela de mortalidade multiestado. Uma tabela de mortalidade multiestado reflete a trajetória de vida de uma coorte à medida que os seus elementos envelhecem e se movimentam entre os estados definidos, assumindo a inexistência de migrações externas (Schoen, 2006d). Estamos, de forma semelhante ao que acontece na tabela de mortalidade no caso simples (secção 1.2), perante uma população estacionária multiestado. Consideremos um modelo multiestado com k estados e um estado adicional, $k + 1$, absorvente (correspondente à morte) e seja $\ell_i(x)$ o número de pessoas no estado i na idade exata x e $d_{ij}(x, x + n)$ o número de transições do estado i para o estado j entre as idades x e

⁶Importa distinguir a função densidade de probabilidade de uma transição de i para j , dada por $\lim_{h \rightarrow 0} \frac{P[X(t)=i, X(t+h)=j]}{h}$, da intensidade de transição, $\mu_{ij}(t)$: a primeira é relativa à probabilidade de transição não condicional, enquanto que $\mu_{ij}(t)$ reflete a densidade de probabilidade de transição condicional à sobrevivência e à ocupação do estado i no início do subintervalo. No modelo linear, a primeira mantém-se constante ao longo do tempo e, no modelo exponencial, é a intensidade de transição que não se altera (Imhoff e Keilman, 1991; Willekens, 2006).

$x + n$. O número de pessoas no estado i na idade exata $x + n$ é igual ao número de pessoas no estado i na idade exata x , menos as saídas para, e mais as entradas de, todos os outros estados do modelo, ou seja:

$$\ell_i(x + n) = \ell_i(x) - \sum_{j \neq i} d_{ij}(x, x + n) + \sum_{i \neq j} d_{ji}(x, x + n)$$

A probabilidade de uma pessoa viva na idade exata x estar no estado i na idade exata $x + n$ é dada por:

$$p_i(x, x + n) = \frac{\ell_i(x + n)}{\ell(x)}$$

Por outro lado, sendo $\ell_{ij}(x + n)$ o número de pessoas no estado j na idade exata $x + n$ que estavam também no estado i na idade exata x , a probabilidade de que uma pessoa no estado i na idade exata x esteja no estado j na idade exata $x + n$ (probabilidade de transição condicional) será:

$$p_{ij}(x, x + n) = \frac{\ell_{ij}(x + n)}{\ell_i(x)}$$

Esta expressão traduz a função probabilidade de transição da cadeia de Markov que reflete o percurso de vida (equação 3) quando a escala de tempo utilizada é a idade (tempo individual).

A taxa de transição do estado i para o estado j entre as idades x e $x + n$, análoga à taxa de mortalidade na tabela de sobrevivência simples, é dada por:

$$m_{ij}(x, x + n) = \frac{d_{ij}(x, x + n)}{L_i(x, x + n)}$$

onde $L_i(x, x + n)$ representa o número de pessoa-anos vividos no estado i entre as idades x e $x + n$, sendo $L_i(x, x + n) = \int_0^n \ell_i(x + u) du$. Já a taxa de transição instantânea do estado i para o estado j na idade exata x (correspondente à equação 4 na secção anterior), pode ser obtida a partir das probabilidades de transição, como:

$$\mu_{ij}(x) = \lim_{u \rightarrow 0} \frac{\pi_{ij}(x, x + u)}{u} \quad i \neq j$$

À semelhança das proporções de sobrevivência apresentadas na equação 2, podemos, com base na tabela de mortalidade multiestado, calcular a proporção de pessoas no estado i entre as idades x e $x + n$ que sobrevivem e estão no estado j exatamente n anos depois, representadas por $s_{ij}(x, x + n)$; esta pode relacionar-se com a probabilidade de transição correspondente, $\pi_{ij}(x, x + n)$. Em termos matriciais, sendo $\Pi(x, x + n)$ a matriz cujo (i, j) -ésimo elemento é $\pi_{ij}(x, x + n)$ e $\mathbf{S}(x, x + n)$ a matriz de proporções de sobrevivência $s_{ij}(x, x + n)$, Hoem e Fong (1976) propuseram a seguinte relação linear:

$$\Pi(x + n/2, x + 3n/2) = \mathbf{S}(x, x + n)$$

1.4.3. População estável multiestado. A população estável multiestado foi introduzida por Rogers (1975) e, analogamente ao discutido na secção 1.3, permite verificar que as implicações a longo prazo de taxas de mortalidade, fecundidade e transições entre estados constantes são uma população com uma composição idade-estado inalterada que cresce exponencialmente a uma taxa intrínseca constante. O recurso a este modelo tem sido moderado pela dificuldade em justificar a manutenção de todas as taxas em valores constantes ao longo do tempo, bem como pelos extensos requisitos de dados necessários à sua aplicação (Schoen, 2006d).

1.4.4. Modelos dinâmicos multi-idades e multiestado. Os modelos multiestado, por considerarem em simultâneo a renovação demográfica e a transição entre diferentes estados no decorrer do tempo de vida, permitem uma análise demográfica mais realista e têm, portanto, um maior poder analítico que os modelos em que somente os nascimentos e as mortes contribuem para alterar a estrutura populacional. Esta vantagem é potenciada quando se permite que as taxas vitais consideradas se alterem ao longo do tempo. O modelo de população estável é baseado em taxas que se mantêm constantes e, como tal, não consegue traduzir as dinâmicas de populações reais. Neste sentido, surge uma generalização da população estável multiestado, para modelos dinâmicos multi-idades e multiestado: o termo multi-idades advém das diferenças na estrutura etária da população que resultam das alterações nas taxas de fecundidade e mortalidade ao longo do tempo; a designação multiestado refere-se ao facto de as alterações nas taxas de transição, que afectam também a estrutura etária, conduzirem, adicionalmente, a diferenças na estrutura por estados. Schoen (2006a) refere modelos em que somente as taxas de fecundidade e mortalidade se alteram ao longo do tempo, mas em que se considera a existência de um único estado em vida, como o modelo metaestável (discreto ou contínuo), os modelos intrinsecamente dinâmicos (do inglês *intrinsically dynamic models*) ou os modelos hiperestáveis. Quando as alterações ocorrem sobre as taxas de transição entre estados, Schoen (2006c) sugere a utilização de modelos multiestado por coorte com taxas de transição proporcionais (adequados ao estudo de coortes, em que a única escala de tempo utilizada é a idade) ou ao recurso a uma abordagem idêntica à dos modelos metaestáveis, intrinsecamente dinâmicos ou hiperestáveis. A integração destes dois tipos principais de modelos, multi-idade e multi-estado, resulta em duas abordagens principais (Schoen, 2006b): o modelo hiperestável (*multi-age and state hyperstable (MASH) model*) e o modelo metaestável (*multi-age and state metastable (MASM) model*). Estes últimos modelos são aqueles que conseguem captar com maior rigor a dinâmica populacional, ao considerarem a possibilidade de alteração simultânea das taxas de fecundidade, mortalidade e transição entre estados ao longo do tempo. Contudo, o seu desenvolvimento encontra-se ainda numa fase inicial, e a sua utilização exige uma extensa informação de base e envolve uma enorme complexidade analítica.

Projeções demográficas

Os objetivos e necessidades específicas dos utilizadores determinam a enorme variedade de projeções demográficas existentes; estas diferem em termos da cobertura geográfica, horizonte temporal ou variáveis demográficas consideradas (O'Neill et al., 2001). Seja qual for o tipo de projeção populacional pretendida, a sua elaboração é habitualmente complexa e deve, portanto, ser o mais estruturada possível. Willekens (1990) delineou um esquema conceptual segundo o qual as projeções demográficas se organizam em oito etapas: (a) identificação do sistema (em que populações, grupos, intervalos de tempo estamos interessados?); (b) descrição do sistema (o que tem acontecido no passado recente?); (c) desenho do modelo (que tipo de modelo de projeção é adequado e exequível?); (d) formulação de hipóteses (como pensamos que se irão alterar a fertilidade, mortalidade e migrações, internas e externas, e sob que pressupostos?); (e) análise de sensibilidade (que pressupostos fazem, de facto, diferença nos resultados?); (f) implementação (de que software dispomos para elaborar as projeções?); e (g) monitorização (como avaliamos as projeções quando o futuro passa a ser passado?). A tomada de decisões, inerente ao cumprimento destas etapas, por parte do investigador, determina em larga escala os resultados das projeções e a sua utilidade.

No presente capítulo, pretende dar-se uma visão geral dos tipos de projeções demográficas habitualmente produzidas, com especial enfoque sobre aquelas que mais se assemelham às projeções alvo deste trabalho, na perspetiva de que tal informação permita facilitar as decisões inerentes ao processo de elaboração das mesmas.

2.1. Principais projeções demográficas produzidas a nível nacional e internacional

As primeiras projeções demográficas de carácter mundial que consideravam explicitamente a estrutura etária e por sexo da população, foram elaboradas por Notestein (1945), do Centro de Investigação sobre População da Universidade de Princeton, E.U.A.. No entanto, muitos esforços no sentido projetar a população ao nível nacional começaram mais de meio século antes (ver, por exemplo, Dorn (1950)). A partir da década de 50, as Nações Unidas conquistaram um papel de liderança na produção de projeções e disseminação dos seus resultados, tendo publicado, entre 1951 e 2008, 21 conjuntos de estimativas (relativas a condições passadas e actuais) e projeções para todos os países e territórios do mundo (O'Neill et al., 2001). Até 1978, estas projeções eram revistas aproximadamente de 5 em 5 anos; desde aí, têm sido elaboradas novas revisões (publicadas na série *World Population Prospects*)

a cada 2 anos. Em intervalos irregulares, as Nações Unidas publicam ainda projeções de longo termo com perspetivas temporais desde 2150 até 2300 (Lutz e KC, 2010).

Outras instituições iniciaram, mais tarde, a elaboração de projeções oficiais, como o Banco Mundial, o *Population Reference Bureau*, o *United States Census Bureau* (USCB), o *International Institute for Applied Systems Analysis* (IIASA) e o Eurostat. O Banco Mundial começou a produzir projeções populacionais em 1978, destinadas sobretudo a uso interno, nomeadamente no âmbito do planeamento para o desenvolvimento, sendo publicadas como parte das séries dos *World Development Reports*. Após 1984, as projeções passaram a ser revistas aproximadamente de 2 em 2 anos sendo, na maioria dos casos, publicada apenas uma variante, mas com um horizonte temporal longo (até 2150). Por volta de 1995, o Banco Mundial deixou de publicar projeções independentes (Lutz e KC, 2010). O *Population Reference Bureau* (PRB) publica projeções da população mundial (apenas para a dimensão populacional e considerando um único cenário), como parte do seu relatório anual. Desde 2000, publicou dimensões populacionais para todos os países e territórios do mundo, projetadas para 2025 e 2050. O *United States Census Bureau* (USCB), por sua vez, produz projeções com base num único cenário para todos os países, desde 1985, com um horizonte temporal variável. Em termos de projeções por região, o *International Institute for Applied Systems Analysis* (IIASA), no âmbito do seu *World Population Programme*, iniciou, em 1994, a produção de projeções globais para 13 regiões do mundo. Especificamente em relação à Europa, as projeções de nível nacional (englobando os Membros da União Europeia e ainda a Noruega e a Suíça) e regional (por regiões de nível NUTS II) produzidas mais recentemente pelo Eurostat são relativas ao período 2008-2060 (Eurostat, 2010).

Em Portugal, o Instituto Nacional de Estatística (INE, I.P. é a entidade oficialmente responsável pela elaboração de projeções demográficas. As mais recentes referem-se ao período 2008-2060 e incidem sobre a população residente em Portugal, por sexos e idades ano a ano até aos 100 anos e mais. A população de base foi a população residente em 1 de janeiro de 2008, de acordo com as “Estimativas Provisórias de População Residente – Portugal, NUTS II, NUTS III e Municípios – 2007”, publicadas pelo INE. (INE, 2009).

Para além das principais organizações cujo campo de acção envolve o estudo da população, investigadores individuais têm contribuído igualmente, de forma significativa, para o desenvolvimento das metodologias de projeção demográfica, tendo-se focado sobretudo na elaboração de projeções ao nível nacional, ou inferior. (O'Neill et al., 2001).

2.2. Técnicas de projeção demográfica

2.2.1. Projeções determinísticas e projeções probabilísticas. As metodologias de projeção demográfica podem ser classificadas em duas categorias principais - determinísticas ou probabilísticas - podendo ainda resultar da integração de técnicas de ambos os grupos. A distinção entre estes dois

tipos principais de metodologias advém da forma de tratar a incerteza associada à previsão do comportamento de todas as componentes da dinâmica populacional, cuja evolução futura é garantidamente incerta.

Nas projeções determinísticas, recorre-se à apresentação de cenários ou variantes que traduzem a combinação de diferentes pressupostos sobre a fecundidade, mortalidade e migrações, de formas consideradas possíveis e consistentes, com o intuito de fornecer uma visão abrangente de como poderá ser o futuro em termos demográficos. Esta abordagem tem sido amplamente utilizada em projeções oficiais: a apresentação de três variantes - elevada, média e baixa - tornou-se comum em projeções elaboradas pelas Nações Unidas, pelo Eurostat, pelo IIASA ou pelo INE, I.P.. Estas variantes permitem considerar futuros alternativos para o desenvolvimento das forças demográficas, mas sem que seja possível quantificar a incerteza associada a cada cenário. Habitualmente, a variante média ou central é percecionada como sendo a que conduz a resultados 'mais prováveis', dado que se baseia no pressuposto de manutenção das tendências verificadas até então em termos das componentes demográficas, mas é necessário ter em conta que: (1) esta maior plausibilidade não pode ser quantificada e (2) esta variante não considera o efeito de políticas futuras já conhecidas ou de outros fatores que possam afectar as tendências demográficas.

No sentido de evitar as limitações associadas às projeções determinísticas, o desenvolvimento de projeções probabilísticas tem sido substancial nas últimas décadas (Wilson e Rees, 2005). Neste tipo de projeções, a incerteza inerente à evolução demográfica é considerada de forma explícita, permitindo associar probabilidades à projeção das diferentes componentes e à projeção da população resultante. Desta forma, obtém-se uma perspectiva mais detalhada da incerteza, podendo ser apresentados intervalos de confiança para traduzir a plausibilidade do futuro demográfico (Skirbekk et al., 2007). A elaboração de projeções probabilísticas passa pela definição de distribuições de incerteza para cada uma das componentes demográficas, que pode ser conduzida através da análise de séries temporais, a partir da análise dos erros de projeções anteriores ou pode ser baseada na avaliação de argumentos de especialistas resultando em conjuntos de distribuições probabilísticas subjetivas (O'Neill et al., 2001). Estas distribuições são depois estocasticamente combinadas num grande número (tipicamente alguns milhares) de projeções, analisando-se a distribuição dos resultados destas simulações. As vantagens deste tipo de projeções são evidentes, mas o seu desenvolvimento encontra-se ainda numa fase inicial e não existe para já uma abordagem universalmente aceite para descrever quantitativamente a incerteza das projeções demográficas, sendo necessária investigação adicional neste campo (Lutz e KC, 2010). A sua utilização é ainda reduzida, continuando a prevalecer as projeções determinísticas. Ao nível global, o IIASA produziu projeções probabilísticas para 13 regiões do mundo (Lutz et al., 1997; Lutz et al., 2001; Lutz et al., 2008); Scherbov e Mamolo (2006) produziram projeções probabilísticas para os então 25 Estados-Membro da União Europeia e, mais tarde, para os 27 que a compõem atualmente (Scherbov et al., 2008), baseando-se nos pressupostos utilizados pelo Eurostat na elaboração de projeções determinísticas e nas respetivas variantes elevada, média e baixa (tendo sido necessário

estabelecer que extensão da incerteza total a ser simulada se encontrava abrangida pela diferença entre as variantes elevada e baixa); a nível nacional, Pedrosa (2011) apresentou recentemente projeções probabilísticas para a população portuguesa, por sexo e idade, até ao ano de 2050.

Os vários tipos de projeções existentes, distinguidos pelas diversas formas de tratar a incerteza, apresentam em comum, na maior parte dos casos, a forma de conjugar a evolução futura dos processos demográficos (fecundidade, mortalidade e migrações) de modo a obter projeções sobre a estrutura e dimensão da população: o método de componentes por coorte é a técnica habitualmente utilizada neste sentido.

2.2.2. Método de componentes por coorte. O desenvolvimento do método de componentes por coorte foi a maior inovação no desenvolvimento das metodologias de projeções demográficas. Esta abordagem foi inicialmente proposta pelo economista inglês Cannan (1895) e formalizada em termos matemáticos por Leslie (1945). A partir da projeção de Notestein (1945), o método das componentes por coorte tornou-se o principal meio de projetar a população e permaneceu fundamentalmente inalterado até à atualidade (O'Neill et al., 2001). Nos dias de hoje, todas as organizações nacionais e internacionais que produzem estatísticas utilizam o método das componentes como ferramenta padrão para a elaboração de projeções demográficas de longo prazo (Lutz e KC, 2010).

De acordo com a descrição de O'Neill et al. (2001), a população inicial do país ou região em estudo é agrupada em coortes definidas, regra geral, pela idade e o sexo, e a projeção processa-se passo a passo através da atualização da população de cada grupo específico de acordo com os pressupostos referentes à fecundidade, mortalidade e migrações que tenham sido estabelecidos. Os grupos etários quinquenais (e os intervalos temporais de cinco anos) são frequentemente utilizados em projeções de longo prazo; como tal, cada coorte sobrevive até ao grupo etário seguinte de acordo com as taxas de mortalidade assumidas para cada período, específicas para a idade e sexo. As migrações podem ser consideradas aplicando-se a cada coorte taxas de migração ou um determinado saldo migratório, também específicos. A dimensão do grupo etário mais jovem é afetada ainda pelo número de nascimentos, calculado com base em taxas de fecundidade relativas às coortes femininas na faixa etária correspondente à idade reprodutiva e, eventualmente, em taxas de mortalidade infantil. É assumida uma determinada relação de masculinidade para repartir os nascimentos totais em masculinos e femininos.

A figura seguinte esquematiza o método das componentes.

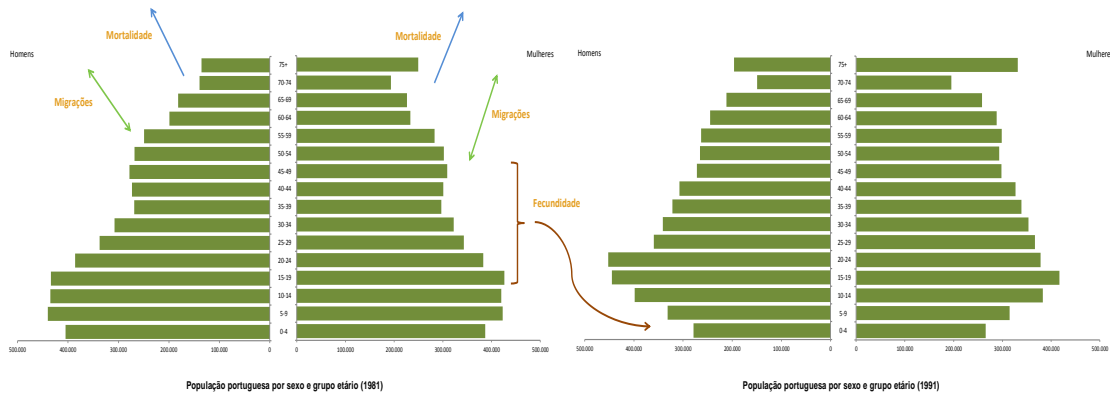


FIGURA 2.2.1. Esquema da projeção demográfica por idade e sexo

Fonte: Adaptado de Lutz et al. (2005)

Uma característica essencial do método é que a dimensão e estrutura etária da população projetada num qualquer momento depende inteiramente da dimensão e estrutura etária iniciais e das taxas de fecundidade, mortalidade e migração assumidas para o período de projeção. A incerteza nos resultados advém, assim, da incerteza sobre a população de partida e sobre as tendências de evolução destas taxas (O'Neill et al., 2001).

O método de componentes por coorte pode ser matematicamente sumarizado através da matriz de Leslie (ver secção 1.3) e de uma equação de projeção que relaciona os efetivos no momento t com a população projetada para o momento $t + 5$ (no caso de projeções quinquenais), considerando ou não a existência de migrações (ver, por exemplo, Siegel (2012)).

A utilização do método de componentes por coorte para projetar uma população por sexo e idade é prática comum e tal metodologia encontra-se bem estabelecida. Contudo, a inclusão de dimensões adicionais a considerar na formação das coortes conduz à necessidade de adaptar o método a uma população caracterizada por um modelo multiestado.

2.2.3. Métodos de projeção demográfica multiestado. Quando se pretendem acrescentar novas dimensões às projeções (considerando atributos adicionais, para além da idade e sexo, para distinguir os indivíduos) pode seguir-se uma das seguintes abordagens principais: o chamado método de distribuição, em que a população projetada por idade e sexo é dividida pelas categorias referentes a outras características, com base em taxas de prevalência ou funções de distribuição (por exemplo, no caso de se pretender dividir a população desagregada por idade e sexo pelos níveis de escolaridade, é comum utilizar-se as taxas de matrícula nos diferentes níveis de ensino); ou o modelo multiestado, baseado na caracterização demográfica através de um modelo populacional multiestado (ver secção 1.4), que leva em conta as probabilidades ou intensidades de transição entre os vários grupos determinados pelas dimensões consideradas. O primeiro é considerado um método estático e, em oposição, o modelo de projeção multiestado, ao focar-se sobre os eventos ocorridos no decorrer de um certo período, para um determinado grupo de indivíduos, incorpora uma perspetiva dinâmica da evolução demográfica.

É considerado, por isso, mais adequado para modelar a evolução demográfica nos casos em que as interações entre grupos desempenham um papel central. No entanto, como discutido anteriormente, a exigência de dados complexos e precisos pode impossibilitar o uso destes modelos e este aspeto conduz ao escasso desenvolvimento de projeções multiestado fiáveis por parte de organismos oficiais (Willekens, 2006). A escolha entre um método estático ou dinâmico tem sido alvo de debate no âmbito da análise e prospetiva demográfica.

2.3. Projeções demográficas por nível de escolaridade

2.3.1. A importância de projetar a população por nível de escolaridade. A escolaridade constitui um elemento essencial da condição humana. O significado original da palavra refere-se ao encaminhamento dos indivíduos (tipicamente jovens) para um estado informado, no qual se pretende que possam produzir apreciações independentes, através da combinação de diferentes capacidades e fragmentos de informação. Este é um importante pré-requisito para superar muitos desafios da vida quotidiana e melhorar a condição humana, tanto ao nível individual como da sociedade (Lutz et al., 2005).

Ao estudar a escolaridade ao nível comunitário, é essencial distinguir entre *stocks* e fluxos. Podemos basear-nos na terminologia adotada por Lutz et al. (2005) para diferenciar a escolaridade, como o processo (fluxo) através do qual as pessoas obtêm competências e conhecimentos, e o *stock* de pessoas escolarizadas, com tais competências e conhecimentos, que constitui o capital humano. De acordo com esta definição, a medição da escolaridade pode incidir sobre dois tipos de variáveis: as que se referem ao fluxo, como, por exemplo, o número de alunos matriculados em cada ano por nível de ensino, e as que se relacionam com o *stock*, como o número de indivíduos por nível de escolaridade mais elevado concluído. As primeiras adequam-se essencialmente ao estudo e avaliação de tópicos que incluem a construção e organização de escolas, formação de professores, a qualidade do ensino ou os conteúdos programáticos. No entanto, para muitas questões do foro social ou económico não interessa saber quantos indivíduos estão a estudar num dado momento, mas antes quantos completaram cada categoria de escolaridade ou qual o número médio de anos de escolaridade completos. Pretende-se, nestes casos, conhecer como se constitui e altera o capital humano que, como se compreende, é altamente influenciado pelo fluxo de escolarização (Lutz e KC, 2010). Existe, porém, um importante hiato temporal na relação entre estas variáveis: o *stock* de capital humano atual resulta do fluxo de escolarização ao longo das décadas precedentes. Por outro lado, dado que a escolaridade é adquirida principalmente na fase inicial da vida, as alterações no capital humano da população adulta refletem-se apenas após um certo intervalo temporal.

A projeção da escolaridade pode ser considerada pertinente se, por um lado, assumirmos que é importante conhecer a composição da população por nível de instrução e, por outro, conseguirmos obter algum benefício significativo a partir do estudo das características futuras desta estrutura. O primeiro

ponto é facilmente fundamentado, já que, como referido anteriormente, a escolaridade é extensamente reconhecida como um impulsor essencial do desenvolvimento, tanto para os indivíduos como para a sociedade. A vasta literatura existente alusiva à relação entre a escolaridade e parâmetros fulcrais no contexto do desenvolvimento, tais como a fecundidade, mortalidade ou crescimento económico (ver, por exemplo, Alachkar e Serow (1988); Haddad et al. (1990); Psacharopoulos e Patrinos (2004)), tem permitido fundamentar os importantes investimentos conduzidos a nível nacional e internacional com vista a promover um maior e melhor acesso à educação. De acordo com o Banco Mundial, a educação é um dos principais instrumentos para promover a redução da pobreza e a melhoria da saúde, a igualdade de género, a paz e a estabilidade (BancoMundial, 2012).

Partindo deste pressuposto devemos, ainda, ponderar sobre as vantagens de tentar projetar para o futuro a evolução da escolaridade e do capital humano. Lutz et al. (2005) distinguem, neste sentido, quatro funções deste tipo de projeções: a) planeamento, b) orientação, c) definição de objetivos realistas, e d) motivação para investimentos de curto prazo que apenas produzem benefícios a longo prazo.

No que respeita ao planeamento governativo no sector da Educação, em que há necessidade de alocar ou gerar recursos (frequentemente escassos), atuando num prazo relativamente curto (5 a 10 anos), a projeção da população de estudantes e professores é muitas vezes relevante, neste caso com maior ênfase do que as alterações na composição educacional da população adulta. Estas modificações futuras no capital humano têm, por sua vez, um papel central na definição de ações para uma orientação estratégica e um planeamento de longo prazo, podendo, inclusivamente, determinar alterações nos balanços geopolíticos de um país. Por outro lado, a definição de objetivos tornou-se um importante elemento das estratégias educativas. Dado que, no curto prazo, os investimentos em educação representam apenas um custo e não produzem um benefício mensurável, as políticas governamentais dispostas a promover a consolidação orçamental podem ser induzidas a cortar nestas despesas educacionais “não produtivas” de modo a reduzir os *deficits* orçamentais existentes. Neste contexto, as projeções da escolaridade que demonstram claramente, de forma quantitativa – e também visualmente, sob a forma das pirâmides etárias educacionais – de que modo é que os aumentos de curto prazo do número de alunos matriculados se traduzem em melhorias de longo termo da composição educativa da população adulta, podem ter um papel fundamental na consciencialização dos governos de que, de facto, o retorno dos seus investimentos chegará, apesar de tal acontecer com uma demora significativa. Em conclusão, parecem existir muitas boas razões para fazer projeções sobre a escolaridade e o capital humano.

2.3.2. A medição do capital humano numa população. De que forma podemos, então, medir o capital humano que pretendemos projetar? Em primeiro lugar, importa salientar que nos iremos focar apenas no estudo da escolaridade formal, dada a dificuldade em obter dados empíricos consistentes que se refiram à aprendizagem ocorrida antes, após ou paralelamente à frequência escolar,

ainda que esta seja também relevante para a acumulação de capital humano. Assim, baseamos o nosso estudo na descrição quantitativa da duração da escolaridade formal. Para tal, podemos recorrer a uma das seguintes variáveis: nível de escolaridade mais elevado completo ou número médio de anos de escolaridade. Este último é um indicador frequentemente utilizado para medir o *stock* de capital humano de uma dada população num determinado ano, e pode ser calculado separadamente para grupos etários distintos. Uma vantagem clara do seu uso é que resulta num único valor, fácil de comparar entre países ou regiões, e que pode ser utilizado de forma simples em análises posteriores (por exemplo, em equações de regressão). No entanto, visto que os dados observados dizem respeito, na maior parte das vezes, ao nível de escolaridade completo, e o cálculo do número médio de anos de escolaridade exige que se façam suposições específicas para o país em estudo (sobre o tempo médio de escolaridade necessário para atingir um dado nível), devendo ainda ser consideradas as taxas de rejeição, o primeiro indicador é muitas vezes escolhido. Considerar a distribuição completa dos níveis de escolaridade atingidos fornece, ainda, informação sobre desigualdades que não é perceptível no caso de se utilizar apenas um valor central, como a média dos anos de escolaridade (Lutz e KC, 2010). O importante debate político sobre a melhoria conjunta dos indicadores relativos à escolaridade primária, secundária e terciária só pode assentar em estudos que considerem explicitamente estas diferentes categorias. Por estas razões, optou-se por estudar a distribuição integral dos níveis de escolaridade completos por sexo e grupo etário.

2.3.3. A relação entre a escolaridade e a dinâmica demográfica. A consciência de que as taxas de mortalidade diferem substancialmente entre grupos socioeconómicos distintos é comum entre os demógrafos (KC et al., 2010), ainda que a existência de uma relação causal entre a educação e a mortalidade não seja, para já, totalmente conhecida (Lleras-Muney, 2005; Clark e Royer, 2010). A nível nacional, tem-se verificado que desigualdades socioeconómicas encontram-se associadas a desigualdades no acesso a cuidados de saúde, nos estilos de vida e nos recursos disponíveis (capital social, cultural, escolar e económico), que se traduzem em diferentes formas de gerir a saúde e a doença ao longo da vida (Fernandes, 2007).

No que se refere ao estudo da relação entre o grau de instrução e a fecundidade, este foca-se essencialmente sobre os indivíduos de sexo feminino: investiga-se a associação entre o nível de instrução das mulheres e o seu efeito sobre a fecundidade feminina. De facto, a escolaridade feminina tem sido identificada, desde há muito, como um dos principais determinantes da fecundidade ao nível individual (KC et al., 2010). A perspetiva mais comumente defendida admite a ideia de que o aumento da escolaridade conduz a uma diminuição da fecundidade, verificando-se, na esmagadora maioria dos casos, que as mulheres mais escolarizadas iniciam a maternidade mais tarde e têm menos filhos no final do período reprodutivo (UN, 1995; Cleland, 2002). Esta relação negativa é explicada sobretudo pelos potenciais efeitos do aumento da escolaridade, nomeadamente no que respeita a uma melhoria do acesso a informação sobre planeamento familiar, a um aumento dos rendimentos e do estatuto social, e a uma maior inclinação para a carreira em detrimento da família. As inter-relações entre estes fatores

conduzem, regra geral, a uma maior empregabilidade feminina e a maiores custos, designadamente custos de oportunidade, associados à maternidade (Oliveira, 2009).

No entanto, de acordo com Jejeebhoy (1995), o impacto da escolaridade feminina sobre a fecundidade pode ser altamente específico do contexto, variando consoante a região do mundo, o nível de desenvolvimento, ou o período em estudo. Nos países em desenvolvimento, o efeito redutor da escolaridade sobre a fecundidade parece ser mais forte do que nos países desenvolvidos. O nível de escolaridade é geralmente mais baixo em países em desenvolvimento e o aumento da escolaridade a partir de níveis baixos, em vez de níveis intermédios, pode ter implicações mais fortes sobre a redução da fecundidade (UN, 1995). Níveis mais baixos de escolaridade podem ser associados a um menor conhecimento sobre reprodução e um menor acesso a métodos contraceptivos. Os indivíduos que têm menores níveis de escolaridade tendem também a ser menos urbanizados, têm provavelmente uma perceção mais tradicional do género e são mais aptos a crer que o estatuto social aumenta através de uma maior fecundidade (Birdsall, 1988; Jejeebhoy, 1995). Tem sido também observado que a ligação a líderes religiosos e a crença de que a prática religiosa requer elevada fecundidade ou proíbe o uso de contraceptivos são mais prevalentes entre os que têm menor escolaridade (McQuillan, 2004). Desta forma, Skirbekk (2008) verificou que o efeito negativo da escolaridade sobre a fecundidade é mais intenso em países da Ásia, África, Médio Oriente e América Latina, relativamente à Europa e América do Norte.

No que diz respeito aos países desenvolvidos, alguns estudos recentes indicam que houve uma transformação na relação entre a fecundidade e a participação das mulheres no mercado de trabalho nos países ocidentais (Brewster, 2000; Sleetbos, 2003; Billari e Philipov, 2004). Investigações baseadas em dados agregados de diferentes países mostram uma associação negativa entre a fecundidade e a empregabilidade feminina até à década de 80, passando a verificar-se uma associação positiva a partir desse período. Atualmente, é nestes países com maior participação feminina no mercado de trabalho que a fecundidade tende a ser mais elevada. Desta forma, será que também se modificou a relação entre a escolaridade e a fecundidade? Relativamente aos países da OCDE, em termos agregados, Sleetbos (2003) assinalou uma correlação positiva entre a escolaridade e a fecundidade durante a década de 90, que se tinha verificado negativa em 1980. Por outro lado, existem alguns estudos baseados em dados individuais, particularmente em Inquéritos sobre a Fecundidade e a Família (*Fertility and Family Surveys*) e investigações similares, em que são igualmente estabelecidos efeitos positivos, mesmo que apenas possam ser verificados em algumas ordens de nascimento (Oliveira, 2009).

Jain (1981) e Gustavsson (2006) sugerem que as condições do mercado de trabalho podem ajudar a explicar a magnitude da relação escolaridade-fecundidade. Esta será mais negativa nos casos em que os custos de oportunidade aumentam com a escolaridade, como acontece quando o emprego e o rendimento se correlacionam positivamente com os níveis de escolaridade. Para além disso, quando existe uma relação negativa entre o número de filhos e a sua escolaridade (num sentido de ‘qualidade

versus quantidade' da descendência), esta pode levar à diminuição dos resultados da fecundidade (Angrist, 2006).

Alguns estudos sugerem que os baixos níveis de fecundidade das mulheres muito escolarizadas, nomeadamente em países caracterizados por níveis de fecundidade abaixo do limiar necessário à substituição de gerações (*below-replacement fertility*), não parecem ser intencionais: os ideais de fecundidade destas mulheres são mais elevados que os das restantes – e, de acordo com alguma evidência, os mais elevados; é a fecundidade real que é baixa (Van Peer, 2002; Testa, 2006).

Oliveira (2009) estudou o caso português e verificou, relativamente ao momento da maternidade e a sua relação com a escolaridade, que: em média, quanto maior o nível de escolaridade da mãe, mais elevada é a sua idade no nascimento do filho. Por outro lado, verificou-se que, na última década, as taxas de fecundidade gerais diminuíram claramente entre os níveis de escolaridade mais baixos e aumentaram ligeiramente entre as mulheres com escolaridade terciária. A diminuição geral da fecundidade em Portugal ter-se-á devido ao decréscimo da fecundidade entre os indivíduos menos escolarizados, tendo sido parcialmente atenuado pelo aumento da fecundidade dos grupos sociais mais elevados. A autora salienta que, no presente, é impossível saber se esta relação não linear se trata de uma nova tendência ou se é um resultado puramente circunstancial, devendo ser discutidos todos os fatores que possam influenciar tal associação. Por um lado, tem existido, nos últimos anos, um crescimento considerável do número de mulheres no mercado de trabalho e um aumento do número de crianças a frequentar educação pré-escolar, o que, também em Portugal, pode conduzir a uma reavaliação da teoria clássica sobre a relação entre a escolaridade e a fecundidade, segundo a qual se admite que existe um conflito entre a empregabilidade da mulher e a maternidade. Por outro lado, estando os efeitos da escolaridade, trabalho e rendimento intrinsecamente ligados, é necessário compreender o efeito da escolaridade no contexto do aumento da prevalência de mulheres empregadas e das diferenças salariais que são consequência do nível de escolaridade.

Até um certo grau de qualificação académica, é possível que o aumento da escolaridade tenha um efeito negativo na fecundidade, por conduzir ao aumento da probabilidade de as mulheres entrarem no mercado de trabalho, embora para posições em que competências educacionais não influenciam os níveis salariais. Para lá desse ponto, a situação muda, e os aumentos no nível de escolaridade resultam em melhores salários; em paralelo, as diferenças na probabilidade de trabalhar não são tão significativas. Assim, maiores salários (no contexto de um acesso generalizado e de uma utilização eficiente dos cuidados infantis, tais como creches) podem resultar numa redução do custo relativo de ter filhos. Consequentemente, as diferenças na fecundidade parecem ser função do nível de escolaridade relacionado com um efeito 'escolaridade-trabalho' entre os grupos menos escolarizados e com um efeito 'escolaridade-rendimento' entre os mais escolarizados (Oliveira, 2009).

Por fim, Dustmann e Glitz (2011) discute a relação entre a escolaridade e as migrações deliberadas⁷, concluindo que as decisões sobre migrações e as decisões sobre investimentos no capital humano se encontram, de facto, estreitamente relacionadas, e que a associação entre ambas de verifica não apenas em relação aos indivíduos que migram, mas também àqueles que permanecem no seu país, seja no país de origem dos migrantes, ou no de destino.

2.3.4. Projeções anteriores. A percepção crescente da importância do capital humano para o desenvolvimento fomentou, nas últimas décadas, várias tentativas para estimar e projetar a composição populacional por nível de escolaridade. Porém, muitos destes esforços não cumpriram o propósito de obter uma matriz completa da estrutura da população por idade, sexo e nível de escolaridade para diferentes pontos no tempo, devido a limitações relacionadas com os dados individuais de cada país ou à ausência de metodologias apropriadas. Psacharopoulos e Arriagada (1992) forneceram estimativas do número médio de anos de escolaridade da população em idade ativa para 99 países. Porém, estas foram obtidas com base em dados censitários e, como tal, apresentam limitações associadas ao facto de se desconhecer o verdadeiro valor do número médio de anos de escolaridade passados em cada nível, dado que as taxas de reprovação e abandono escolar teriam de ser estimadas pelos investigadores. Nehru et al. (1995) estimaram o número médio de anos de escolaridade nos níveis primário, secundário e terciário para a população em idade ativa, mas sem desagregar as estimativas por sexo. Barro e Lee (1993); Barro e Lee (2000) produziram dados sobre os níveis de escolaridade completos e o número médio de anos de escolaridade para um vasto conjunto de países, também a partir de dados censitários. No entanto, as estimativas obtidas referiam-se a apenas dois amplos grupos etários (15 e mais anos e 25 e mais anos), e somente para o período 1960-2000. Ahuja e Filmer (1995a) partiram de projeções demográficas da Organização Internacional do Trabalho por sexo e grupo etário e sobrepueram-lhes uma distribuição estimada da escolaridade para dois amplos grupos etários (6-24 e 25 e mais anos), obtida a partir de dados sobre as proporções de alunos matriculados e respetivas projeções fornecidas pela UNESCO. À semelhança de Nehru et al. (1995), utilizam o chamado Método do Inventário Permanente, em que se procede à soma do total de alunos matriculados no decorrer de longas séries temporais, obtendo-se depois estimativas da obtenção de escolaridade na população adulta (proporção da população por nível de escolaridade completo e/ou número médio de anos de escolaridade). Visto que estas longas séries temporais estão raramente disponíveis, este método envolve muitas projeções para o passado baseadas num conjunto de pressupostos. Para além disso, podemos dizer que esta metodologia apresenta, de alguma forma, uma natureza estática, por não permitir que a composição da população em termos de escolaridade influencie a fecundidade ou a mortalidade (Lutz et al., 2005).

A aplicação do modelo multiestado à projeção por sexo, idade e nível de escolaridade tem sido conduzida essencialmente no IIASA. Vários membros deste instituto austríaco têm elaborado projeções por

⁷Dustmann e Glitz (2011) distingue dois tipos de migrações: as que resultam de movimentações forçadas por desastres naturais ou circunstâncias criadas pelo homem (como perseguições devidas a visões políticas ou religiosas distintas), e as que se devem a decisões individuais, baseadas na busca de condições económicas mais favoráveis noutras regiões, sendo estas classificadas como migrações deliberadas.

nível de escolaridade para uma série de países (Goujon e Wils, 1996; Lutz, 1999; Lutz e Goujon, 2001; Lutz et al., 2005; Goujon et al., 2007; KC et al., 2010), tentando, fundamentalmente, relacioná-las com o seu nível de desenvolvimento. A abordagem mais próxima à seguida pelo IIASA é a do *Education Policy and Data Center* (EPDC), cujo modelo (EDPOP) foi desenvolvido por Annababette Wils. Em 2007, o EPDC produziu projeções para 83 países em desenvolvimento e três categorias de nível de escolaridade até 2025, baseando-se na definição de trajetórias específicas por país (Wils, 2007). Nestas projeções, tem sido aplicada uma amordagem dinâmica, procurando evitar parte das limitações associadas aos trabalhos anteriores. Recentemente, o IIASA desenvolveu, em colaboração com o Eurostat e o Instituto Nacional de Estatística do Reino Unido (*UK Office of National Statistics*), uma nova abordagem para avaliar argumentos sobre hipóteses alternativas relativas a possíveis tendências futuras, com base num grande número de questionários, determinando a avaliação dos argumentos alternativos por especialistas. Esta abordagem está a ser atualmente convertida num novo conjunto de projeções para todos os países do mundo por idade, sexo e nível de escolaridade, com publicação prevista para 2012⁸ (Lutz e KC, 2010).

⁸Esta publicação não ocorreu até ao momento de entrega da presente dissertação.

Caracterização demográfica recente da população portuguesa

A estrutura demográfica da população portuguesa tem vindo a modificar-se de forma evidente nas últimas décadas, encontrando-se tais mudanças estreitamente relacionadas com transformações de carácter político, económico e social, num encadeamento de influências recíprocas (Carrilho e Patrício, 2010). A evolução de cada uma das principais componentes demográficas – mortalidade, fecundidade e migrações – espelha os efeitos, em muitos casos perceptíveis apenas no longo prazo, de um conjunto de acontecimentos ou medidas que determinaram, de forma mais ou menos direta, os comportamentos de cada indivíduo que constitui a população. Ao mesmo tempo, as mudanças demográficas verificadas têm dado origem, elas mesmas, a decisões no sentido de adequar as políticas e práticas governativas à composição da população ou de conduzir a alterações demográficas favoráveis, de acordo com a estratégia de governação do país (Barreto, 2004). O balanço conjunto da evolução destas componentes determina a dimensão e a estrutura populacional em cada momento.

3.1. Componentes demográficas

3.1.1. Mortalidade. A melhoria do acesso aos cuidados de saúde, o aumento significativo do conhecimento, em particular no campo da medicina, e os avanços tecnológicos que permitiram desenvolver substancialmente as técnicas de diagnóstico e tratamento, contribuíram, juntamente com fatores de ordem económica e social que traduzem uma melhoria na qualidade de vida da população portuguesa nas últimas décadas, para aumentar consideravelmente o número médio de anos que podemos esperar viver (Carrilho e Patrício, 2010). A esperança média de vida, com referência tanto ao momento do nascimento como a idades posteriores, situa-se em níveis cada vez mais elevados (figuras 3.1.1 e 3.1.2). Contudo, as diferenças associadas ao género são evidentes, sendo que as mulheres apresentam, regra geral, vantagens em termos de sobrevivência: em termos biológicos, as mulheres parecem apresentar maior resistência no que respeita às doenças do aparelho circulatório e ao processo de envelhecimento; para além disso, tendem a estabelecer relações sociais mais intensas e com um maior número de pessoas, o que propicia a que recebam mais apoio e consigam adaptar-se com maior facilidade a situações traumáticas que ocorrem ao longo da vida (por exemplo, a situações de doença ou viuvez); por outro lado, os homens mantêm-se mais expostos a riscos associados a uma mortalidade prematura, já que tendem a exercer profissões de maior risco para a saúde, a apresentar um consumo de tabaco e álcool mais elevado e a conduzir de modo menos prudente; por fim, as condições de saúde reprodutiva das mulheres melhoraram de forma significativa, com a diminuição

substancial da mortalidade associada à gravidez e ao parto, permitindo que a influência destes fatores sobre a sobrevivência das mulheres diminuísse drasticamente (Fernandes, 2007; Oliveira e Mendes, 2010).

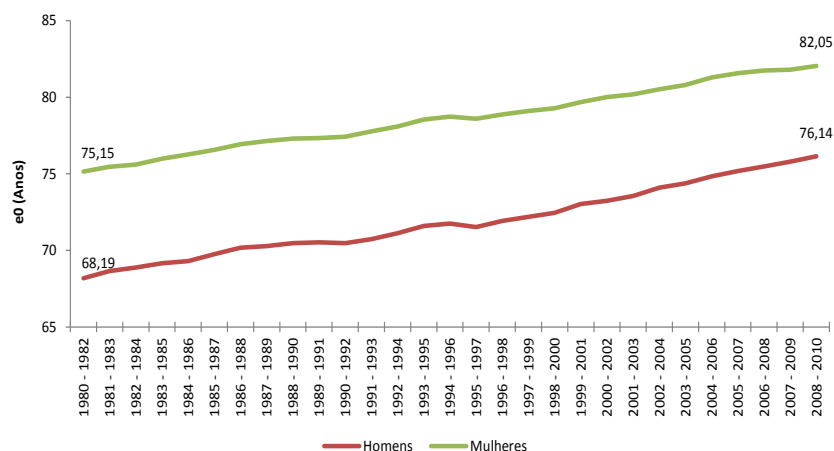


FIGURA 3.1.1. *Esperança de vida à nascença, por sexo, 1980-2010*

Fonte: INE, *Tábuas Completas de Mortalidade*

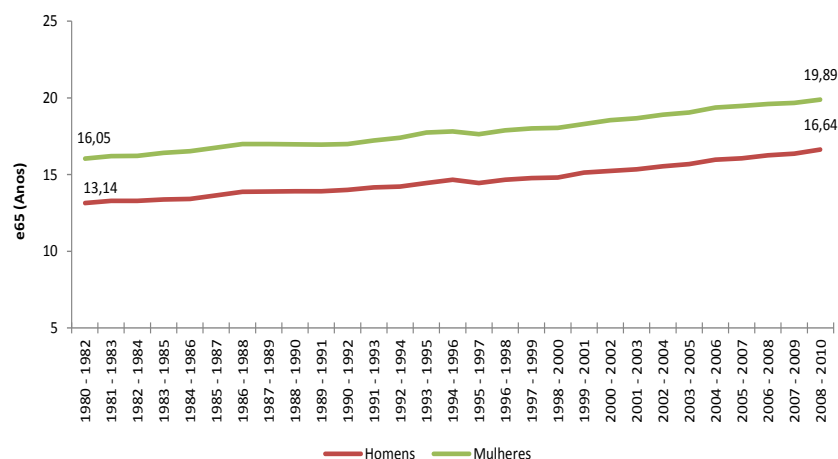


FIGURA 3.1.2. *Esperança de vida aos 65 anos, por sexo, 1980-2010*

Fonte: INE, *Tábuas Completas de Mortalidade*

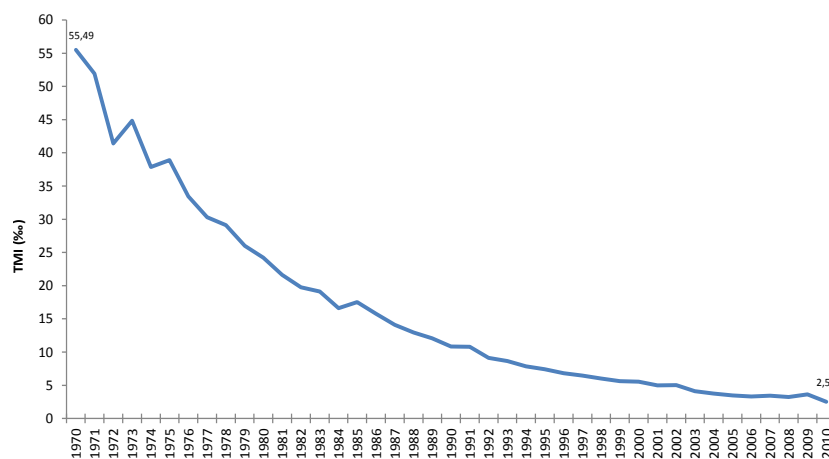
Pelas mesmas razões, as diferenças entre os géneros verificam-se também quando analisamos a evolução da taxa bruta de mortalidade (tabela 3.1.1). O número de óbitos por 1000 habitantes tem apresentado ligeiras variações ao longo do tempo, sendo sempre mais elevado entre os indivíduos do sexo masculino.

A taxa de mortalidade infantil registou um decréscimo acelerado nas últimas décadas (figura 3.1.3): entre 1970 e 2009, o número de óbitos nos primeiros 12 meses de vida, por mil nados-vivos, diminuiu, em média, 6,8% ao ano, correspondendo à taxa de declínio mais elevada entre os países da OCDE (OCDE, 2011) e, em 2010, este indicador situou-se nos 2,53‰ (dados provisórios), muito próximo de países como a Suécia (2,54‰), Finlândia (2,3‰) e Islândia (2,2‰) (Eurostat, 2012). Tais resultados

TABELA 3.1.1. *Taxa Bruta de Mortalidade, por sexo, 1970-2010*

Fonte: INE, Indicadores Demográficos

Ano	Taxa Bruta de Mortalidade (‰)		
	HM	H	M
1970	10,72	11,46	10,06
1980	9,72	10,55	8,96
1990	10,29	11,05	9,59
2000	10,3	11,15	9,51
2010	9,96	10,53	9,42

FIGURA 3.1.3. *Taxa de mortalidade infantil, 1970-2010*

Fonte: INE, Indicadores Demográficos

refletem a melhoria dos cuidados de saúde materno-infantis, com a generalização do planeamento familiar e com destaque para o Programa Nacional de Vacinação (MS, 2009).

A análise comparativa dos níveis de mortalidade entre indivíduos com diferentes níveis de escolaridade⁹ é dificultada pela inexistência de um sistema de registo completo que permita conhecer o nível de instrução de cada pessoa falecida. Apenas são disponibilizados pelo INE, I.P., dados anuais relativos ao número de óbitos com menos de 1 ano e ao número de nados-vivos por nível de escolaridade mais elevado completo da mãe; a partir destes, podemos calcular as taxas de mortalidade infantil por nível de instrução da mãe. A análise da evolução destas taxas para o período 1996-2010 (figura 3.1.4) permite verificar que o decréscimo da mortalidade infantil tem ocorrido em todos os grupos de escolaridade da mãe, existindo um diferencial bastante evidente: no período estudado, a taxa de mortalidade infantil no grupo de mães sem escolaridade foi, em média, 2,8 vezes superior à taxa de mortalidade infantil no grupo de mães que tinham o Ensino Superior completo, no mesmo ano. Os dados obtidos sugerem, assim, que quanto mais elevada é a escolaridade da mãe, menor o nível de mortalidade infantil associado. Este efeito parece revelar-se sobretudo a partir da conclusão do Ensino Secundário, já que as taxas correspondentes às mães sem escolaridade e às que completaram o Ensino Básico têm vindo a apresentar comportamentos próximos. De facto, o teste de Friedman

⁹A definição dos grupos de escolaridade utilizados é discutida somente na secção 5.2.1; contudo, opta-se por utilizar desde já a mesma classificação, com vista a manter a coerência do trabalho.

aponta para a existência de diferenças estatisticamente significativas nas taxas de mortalidade infantil observadas nos diversos anos, por nível de escolaridade da mãe, para qualquer nível de significância habitualmente utilizado ($p < 0.001$). Analisando os grupos de escolaridade da mãe par a par, através do teste de Wilcoxon (*signed-ranks*) para amostras emparelhadas, pode admitir-se, para um nível de significância de 5%, que a mediana das diferenças é estatisticamente diferente de zero, seja qual for o par considerado ($p = 0.031$ para as diferenças entre as taxas correspondentes a mães sem escolaridade e com Ensino Básico e $p = 0.001$ para as restantes combinações).

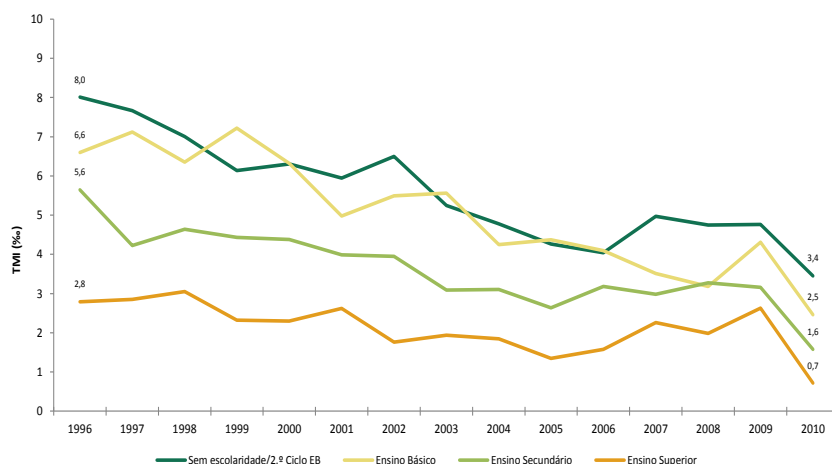


FIGURA 3.1.4. Taxa de mortalidade infantil por nível de escolaridade da mãe, 1996-2010

Fonte: Cálculos do autor

3.1.2. Natalidade e Fecundidade. O número de nados-vivos por 1000 habitantes desceu consideravelmente desde a década de 70. Manteve-se em torno dos 11 nascimentos durante a década de 90 e, após um ligeiro aumento entre os anos de 1995 e 2000, tem vindo a diminuir desde esse momento. Em 2010 foram registados 9,5 nados-vivos por 1000 habitantes (tabela 3.1.2).

TABELA 3.1.2. Taxa Bruta de Natalidade, 1970-2010

* Quebra de série

Fonte: PORDATA; INE, Indicadores Demográficos

Ano	Taxa Bruta de Natalidade (‰)
1970	20,8
1980	16,2*
1990	11,7
2000	11,7
2010	9,5

De facto, o padrão de fecundidade em Portugal tem vindo a caracterizar-se, nas últimas décadas, por uma tendência evidente de diminuição do número médio de filhos por mulher em idade fértil (figura 3.1.5): em 1980, este valor situava-se em 2,8 filhos e, em 2010, era de apenas 1,37 filhos por mulher¹⁰. Como refere Oliveira (2008), a descida dos níveis de fecundidade insere-se num conjunto de

¹⁰Desde o ano de 1982 que o número médio de filhos por mulher em idade fértil (Índice Sintético de Fecundidade) se situa abaixo daquele que é considerado o nível mínimo de substituição de gerações nos países mais desenvolvidos (INE, 2012), correspondente a 2,1 crianças por mulher.

alterações sociais e demográficas caracterizadas pela manutenção da fecundidade abaixo do nível de substituição das gerações, pelo adiamento do casamento e dos nascimentos, pelo aumento do divórcio, da coabitação sem casamento formal e dos filhos fora do casamento. A procriação ocorre em idades cada vez mais tardias, tendo vindo a idade média da mãe ao nascimento do primeiro filho a aumentar gradualmente desde a década de 80, aproximando-se atualmente dos 29 anos.

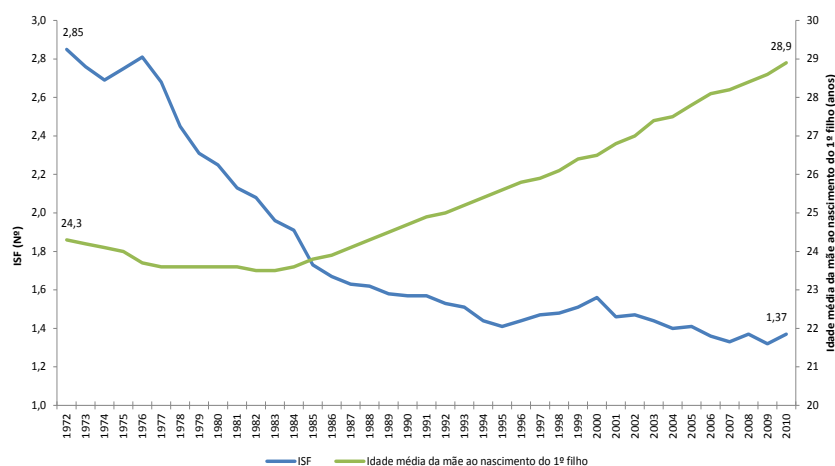


FIGURA 3.1.5. *Índice Sintético de Fecundidade e idade média da mãe ao nascimento do primeiro filho, 1972-2010*

Fonte: INE, Indicadores Demográficos

As diferenças no comportamento reprodutivo por nível de escolaridade evidenciam-se principalmente no que diz respeito ao momento em que ocorre a fecundidade. As mulheres com níveis de instrução mais elevados tendem a ter os seus filhos mais tarde, o que se relaciona com a intenção habitual de concluir o percurso escolar antes da maternidade. Com base nos dados relativos ao ano de 2010 (figura 3.1.6), verificamos que as mulheres sem escolaridade formal têm os seus filhos sobretudo antes dos 30 anos, idade em que aumenta consideravelmente o número de nascimentos entre aquelas que concluíram o Ensino Superior.

A evolução do número médio de filhos por mulher por nível de escolaridade para o período 1998-2010 (figura 3.1.7) sugere que: nos últimos anos, as mulheres que concluíram o Ensino Básico são aquelas que têm, em média, menos filhos; para além disso, se por volta do ano 2000 eram as mulheres com o Ensino Superior que tinham um maior número de filhos, atualmente são as que não têm escolaridade formal que lideram neste indicador, com uma média de 1,62 filhos por mulher (de facto, foi apenas no grupo das mulheres sem escolaridade que este valor foi superior em 2010 relativamente ao início do período estudado). O teste de Friedman sugere que estas diferenças no ISF por nível de escolaridade da mãe observadas ao longo do tempo são estatisticamente significativas ($p < 0.001$). Os testes de Wilcoxon (*signed-ranks*) para cada par de grupos de escolaridade da mãe apontam para a existência de diferenças na distribuição do ISF entre as mães com o Ensino Básico e todos os outros grupos ($p = 0.003$, $p = 0.002$ e $p = 0.001$ para as diferenças entre este grupo e as mães sem escolaridade, as

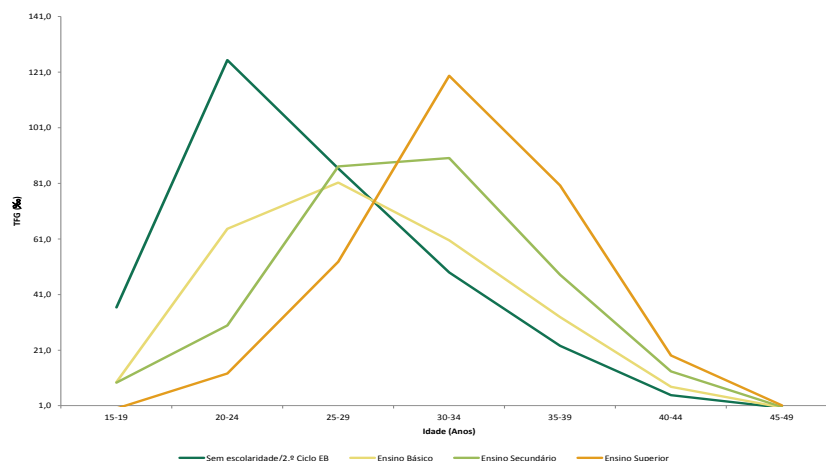


FIGURA 3.1.6. Taxa de Fecundidade, por grupo etário e nível de escolaridade da mãe, 2010

Fonte: Cálculos do autor

que completaram o Ensino Secundário e as que completaram o Ensino Superior, respetivamente), e entre as mães com o Ensino Secundário e o Ensino Superior ($p = 0.012$).



FIGURA 3.1.7. Índice Sintético de Fecundidade, por nível de escolaridade da mãe, 1998-2010

Fonte: Cálculos do autor

3.1.3. Migrações. A evolução das migrações nos anos mais recentes caracteriza-se pelo importante contributo do saldo migratório para o crescimento populacional em Portugal, particularmente desde o início da década de 90 (figura 3.1.8). A partir de 1994, a diferença entre o número de imigrantes e de emigrantes tem ultrapassado consideravelmente o balanço entre o número de nascimentos e de mortes (saldo natural). Apesar de o saldo migratório se manter positivo desde 1993, verifica-se que, ao contrário do que aconteceu até 2002, atualmente o número de entradas aproxima-se cada vez mais do número de saídas: em 2010, esta diferença foi de apenas 3815 indivíduos, quando em 2002 se situava em 70000 pessoas. Como tal, o contributo da componente migratória para o crescimento populacional diminuiu, deixando de ser suficiente para garantir o aumento do número de residentes

em Portugal, também perante a diminuição do saldo natural, que se situa, desde 2007, em valores negativos.

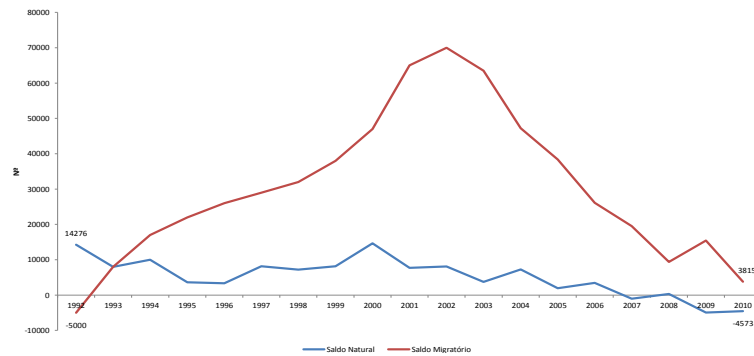


FIGURA 3.1.8. *Saldo Natural e Saldo Migratório, 1992-2010*

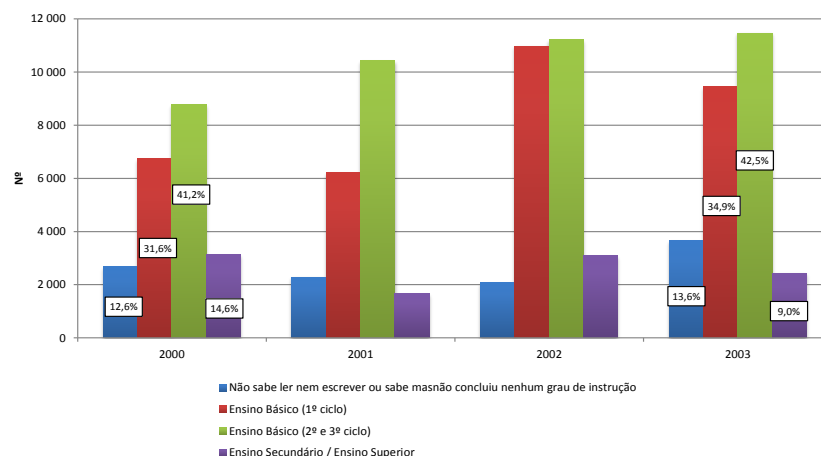
Fonte: INE, Indicadores Demográficos

Para além de representarem, nos dias de hoje, o fator com maior influência sobre a variação do número de indivíduos nas populações dos países desenvolvidos (dada a evolução do saldo natural, cuja contribuição neste sentido é cada vez menor), as migrações têm, geralmente num longo prazo, um relevante impacto nos comportamentos demográficos relacionados com a fecundidade e com a mortalidade; de modo mais imediato, contribuem também para alterar a estrutura populacional, tanto em relação ao sexo como à estrutura etária. De acordo com Carrilho e Patrício (2010), a população imigrante é mais jovem que a população nacional e concentra-se na faixa da população em idade ativa, podendo colmatar o efeito da emigração, que afecta sobretudo este grupo etário; por outro lado, a maioria dos migrantes são do sexo masculino, ao contrário da distribuição da população nacional, em que as mulheres constituem o grupo maioritário.

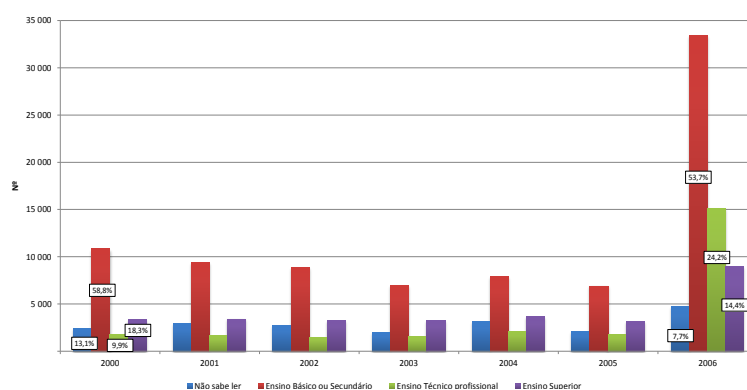
A escassez de dados relativos às diferenças no padrão migratório por nível de escolaridade dificulta o estudo do efeito das migrações sobre a distribuição da população residente de acordo com o nível de instrução. Os dados disponibilizados pelo INE, I.P., apesar de relativos a curtos períodos temporais, sugerem que grande parte dos emigrantes e dos indivíduos estrangeiros que solicitam estatuto de residente possuem um nível de instrução intermédio (Ensino Básico ou Ensino Secundário); adicionalmente, entre os imigrantes parece existir sempre uma maior percentagem de indivíduos com escolaridade de nível superior relativamente àqueles que não têm escolaridade formal, enquanto que, entre os emigrantes, este padrão já não se verifica (figuras 3.1.9 e 3.1.10).

3.2. Fatores de crescimento e grupos funcionais

A diminuição dos níveis de fecundidade tem sido determinante para o decréscimo da taxa de crescimento natural, que adquire valores negativos desde 2007 (com valor nulo em 2008); por outro lado,

FIGURA 3.1.9. *Emigrantes, por nível de instrução, 2000-2003*

Fonte: INE, Estatísticas Demográficas

FIGURA 3.1.10. *População estrangeira que solicitou estatuto de residente, por nível de instrução, 2000-2006*

Fonte: INE, Estatísticas Demográficas

como discutido anteriormente, o balanço entre as entradas e saídas de indivíduos detém um peso cada vez menor sobre o aumento do número de residentes em Portugal: em 2010, o crescimento migratório não foi suficiente para compensar o valor negativo de crescimento natural e, tal como ocorria no final da década de 80 e início da década de 90, o crescimento efetivo da população voltou a ser negativo (figura 3.2.1).

O aumento da esperança média de vida tem conduzido a uma maior representatividade dos indivíduos com mais de 65 anos na população: entre 1992 e 2010, o número de idosos¹¹ subiu 36,7%, e correspondia, no final deste período, a 18% da população (figura 3.2.2). Paralelamente, o decréscimo das taxas de fecundidade traduz-se numa diminuição progressiva da percentagem de jovens (indivíduos com idades até 14 anos) que, desde o ano 2000, tem sido inferior à proporção de idosos. A percentagem de indivíduos potencialmente activos tem-se mantido aproximadamente constante em torno dos 67%, mas devemos ter em conta que os efeitos da alteração dos padrões de fecundidade e mortalidade se

¹¹Utiliza-se a classificação "idoso" para designar os indivíduos com 65 ou mais anos, sendo esta a idade cronológica aceite na maioria dos países desenvolvidos para definir este grupo (WHO, 2012).

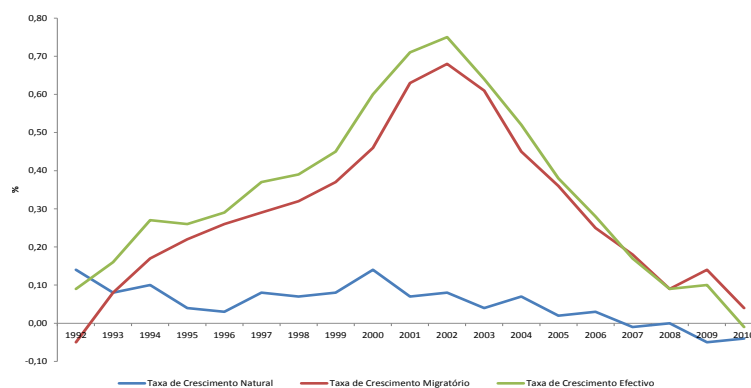


FIGURA 3.2.1. Taxas de Crescimento Natural, Migratório e Efetivo, 1992-2010

Fonte: INE, Indicadores Demográficos

deverão fazer sentir também sobre este grupo, apesar de, como é óbvio, actuarem num prazo mais longo. Neste sentido, podemos esperar que a representatividade dos indivíduos entre os 15 e 64 anos venha também a diminuir, com o aumento continuado da proporção de idosos.

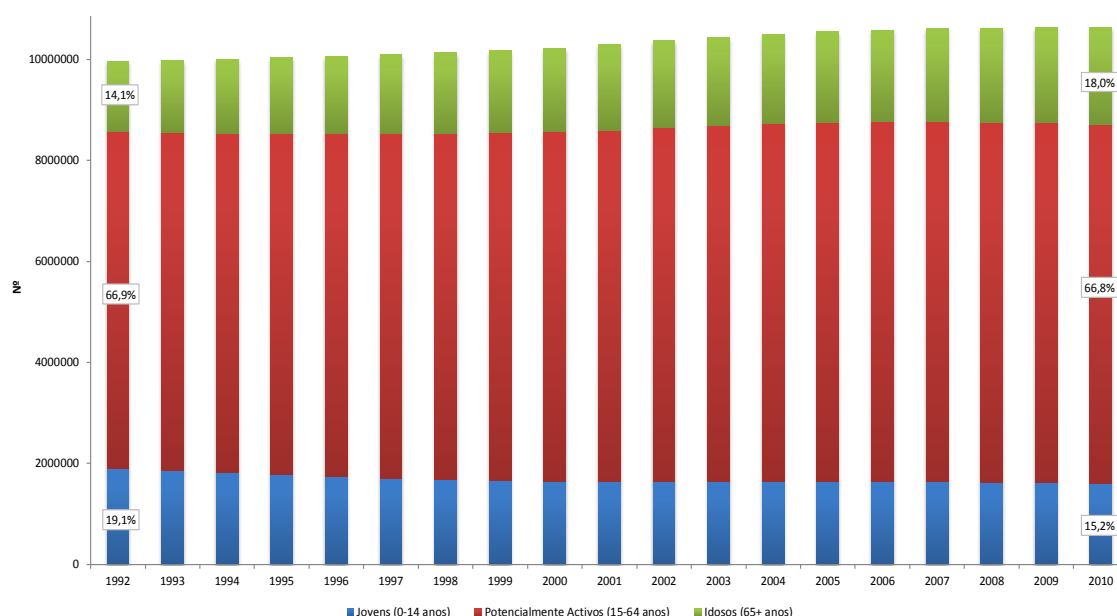


FIGURA 3.2.2. Grupos Funcionais, 1992-2010

Fonte: INE, Estimativas Anuais da População Residente

Tais alterações refletem-se igualmente na evolução dos índices resumo (figura 3.2.3): o aumento do índice de envelhecimento comprova o acréscimo do número de idosos relativamente ao de jovens (em 1971, existiam aproximadamente 34 idosos por cada 100 jovens e, em 2010, este valor ascendia já a 120 idosos por 100 jovens). Por outro lado, o peso dos idosos mais velhos (com 75 ou mais anos) na população idosa total tem vindo a agravar-se (a percentagem de indivíduos com 75 ou mais anos entre o total de idosos, correspondente ao índice de longevidade, chegou em 2010 aos 47,4%). No que se refere à estrutura etária da população em idade ativa, denota-se que a mesma tem vindo a envelhecer, já que, por 100 pessoas entre os 55 e os 64 anos, existiam em 2010 aproximadamente 103 pessoas com

20 a 29 anos, quando o mesmo valor em 1971 era de 142,4 pessoas. Dado que se pode esperar que a tendência de envelhecimento se mantenha, a franja mais velha da população em idade ativa deverá exceder brevemente o número de pessoas entre os 20 e os 29 anos.

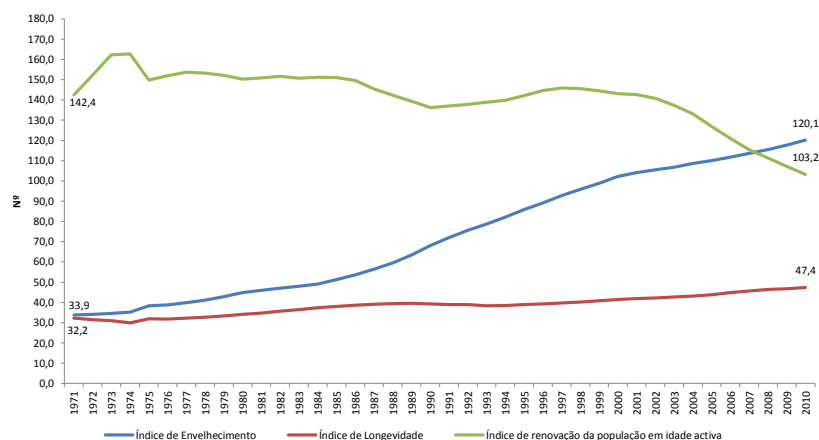


FIGURA 3.2.3. Índices de envelhecimento, longevidade e renovação da população em idade ativa, 1971-2010

Fonte: INE, Estimativas Anuais da População Residente e Indicadores Demográficos

Por outro lado, o número de idosos por cada 100 pessoas em idade ativa (índice de dependência de idosos), aumentou de 15,66 em 1971, para 27,2 em 2010 (figura 3.2.4). À medida que o número de jovens em relação à população em idade ativa diminui, verifica-se que o índice de dependência total, que reflete o balanço entre estes dois indicadores, tendo diminuído até à década de 90, tem vindo a aumentar desde então.

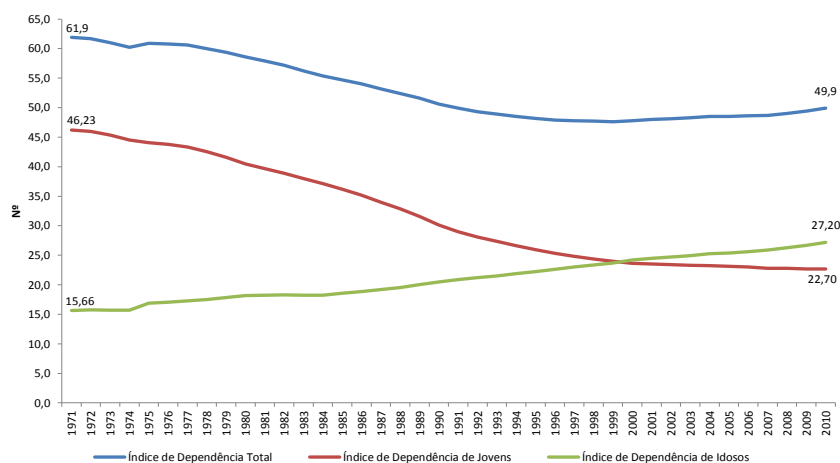


FIGURA 3.2.4. Índices de dependência total, de jovens e de idosos, 1971-2010

Fonte: INE, Estimativas Anuais da População Residente

Estes desenvolvimentos representam importantes desafios no âmbito das decisões de governação, dado o seu impacto inevitável no sector das políticas públicas, nomeadamente em relação aos apoios sociais, aos cuidados de saúde ou ao mercado de trabalho. A projecção para o futuro das tendências verificadas é, como tal, de enorme relevância.

A Educação em Portugal

Para a elaboração de projeções por nível de escolaridade, e independentemente da metodologia utilizada, é evidente a necessidade de conhecer as especificidades do Sistema Educativo pelo qual se rege a população em estudo, bem como a tendência relativa à escolarização verificada no passado, de forma a garantir que os pressupostos admitidos se adequam à realidade.

4.1. A estrutura atual do Sistema Educativo Português e reformas em curso

O Sistema Educativo Português encontra-se organizado em quatro níveis principais (figura 4.1.1): Educação Pré-escolar, Ensino Básico, Ensino Secundário e Ensino Superior (GEPE, 2012).

A Educação Pré-escolar, de frequência facultativa, destina-se a crianças a partir dos 3 anos de idade e estende-se até à entrada no Ensino Básico que ocorre, habitualmente, aos 6 anos. Nessa altura, as crianças iniciam a escolaridade obrigatória, estruturada em três níveis e com a duração total de 9 anos: até aos 10 anos, as crianças devem frequentar o 1.º Ciclo do Ensino Básico, que corresponde aos primeiros quatro anos de escolaridade obrigatória; o 2.º Ciclo prolonga-se por dois anos e, regra geral, aos 12 anos de idade, dá-se a entrada no 3.º Ciclo. Os alunos que tenham concluído a escolaridade obrigatória ou possuam habilitação equivalente poderão depois frequentar o Ensino Secundário. Os cursos de nível secundário, com a duração de três anos, podem ser, atualmente, de quatro tipos: cursos científico-humanísticos, orientados fundamentalmente para o prosseguimento de estudos de nível superior; cursos tecnológicos, direcionados a alunos que pretendem entrar no mercado de trabalho, possibilitando ainda a prossecução de estudos em cursos tecnológicos especializados ou no Ensino Superior; cursos artísticos especializados, vocacionados a assegurar formação artística especializada nas áreas de artes visuais, audiovisuais, dança e música, permitindo a entrada no mercado de trabalho ou a continuação de estudos em cursos pós-secundários não superiores ou no Ensino Superior; e cursos profissionais, organizados por módulos em diferentes áreas de formação, que também possibilitam tanto a inserção no mercado de trabalho, como o prosseguimento de estudos em cursos pós-secundários não superiores ou no Ensino Superior (GEPE, 2012).

Nos últimos anos, têm vindo a ser promovidas algumas modificações a este esquema de organização (ME, 2012). A Lei n.º 85/2009, de 27 de Agosto, veio determinar o regime da escolaridade obrigatória para as crianças e jovens entre os 6 e os 18 anos de idade (considerados em idade escolar) e estabelecer a universalidade da educação pré-escolar para as crianças a partir dos 5 anos de idade. Em sequência, ao

pretender-se assegurar que todos os alunos frequentem estabelecimentos de educação ou de formação pelo menos entre os 5 e os 18 anos de idade, torna-se necessário proceder ao reordenamento da rede escolar, cujos critérios foram definidos pela Resolução do Conselho de Ministros n.º 44/2010, de modo a: adaptar a rede escolar ao objetivo de uma escolaridade de 12 anos para todos os alunos; adequar a dimensão e as condições das escolas à promoção do sucesso escolar e ao combate ao abandono e promover a racionalização dos agrupamentos de escolas. Por outro lado, o Despacho Normativo n.º 1/2005, de 5 de janeiro, prevê a possibilidade de, em casos excecionais, se acelerar a progressão no Ensino Básico, podendo o 1.º Ciclo ser completado em três anos. Complementarmente, o Despacho normativo n.º 29/2010 estabelece que a transição ao 2.º Ciclo do Ensino Básico pode ocorrer aos 8 anos de idade, para alunos que tenham beneficiado do regime de antecipação da matrícula no 1.º ano de escolaridade.

Após a conclusão do Ensino Secundário, habitualmente por volta dos 18 anos, os alunos podem ingressar no Ensino Pós-secundário não superior, através dos chamados cursos de especialização tecnológica (CET). Estes cursos visam conceder formação especializada em diferentes áreas tecnológicas e permitem a entrada no mercado de trabalho ou o seguimento de estudos de nível superior (GEPE, 2012).

Os indivíduos que tenham interrompido precocemente o seu percurso escolar ou estejam em risco de o fazer, aqueles que não puderam frequentar a escola durante a juventude e quem pretender prosseguir a sua formação ao longo da vida, encontra a possibilidade de regressar à escola através da educação e formação de jovens e adultos. Este tipo de formação abrange seis diferentes modalidades, que podem atribuir uma certificação escolar e/ou uma qualificação profissional, possibilitando ainda a continuação de estudos de nível pós-secundário não superior ou o Ensino Superior (GEPE, 2012).

O ingresso em cada instituição de Ensino Superior está sujeito a *numerus clausus*, podendo candidatar-se ao acesso a este nível de ensino os alunos que: tenham concluído um curso de ensino secundário ou uma qualificação legalmente equivalente; tenham realizado os necessários exames de admissão para o curso que desejam frequentar com uma classificação mínima de 95 pontos e, quando aplicável, satisfaçam os pré-requisitos exigidos para o curso a que se candidatam. Para além disso, também os alunos com idade superior a 23 anos que não possuem habilitações para o ensino superior podem tentar aceder através da realização de exames específicos, organizados pelos respetivos estabelecimentos de ensino superior, que demonstrem a sua capacidade para frequentar o curso a que se candidatam. O Ensino Superior organiza-se, atualmente, segundo os princípios de Bolonha. São conferidas as seguintes qualificações académicas: o grau de Licenciado, cujo ciclo de estudos pode ter a duração de 6 ou até 8 semestres, conforme seja atribuído por um Instituto Politécnico ou por uma Universidade, respetivamente; o grau de Mestre, com um ciclo de estudos correspondente a 3 ou 4 semestres e o grau de Doutor, atribuído a quem tenha obtido aprovação nas unidades curriculares do curso de Doutoramento, quando exista, e no ato público de defesa da tese (GEPE, 2012).

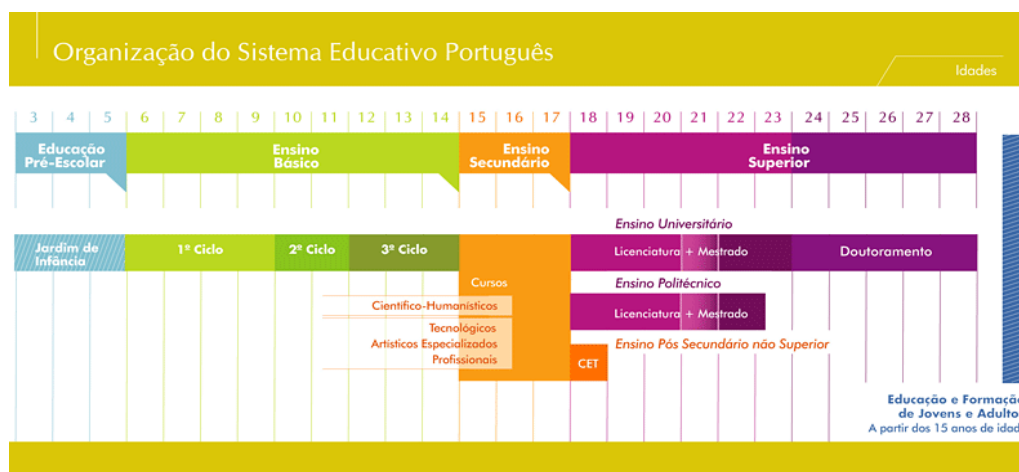


FIGURA 4.1.1. Organização do Sistema Educativo Português

Fonte: Gabinete de Estatística e Planeamento da Educação, Ministério da Educação

4.2. Indicadores de escolarização da população portuguesa

4.2.1. Taxa Real de Escolarização, aprendizagem ao longo da vida, abandono precoce de educação e formação, inscritos e diplomados no Ensino Superior. Considerando os níveis de ensino até ao Ensino Secundário, verifica-se que a percentagem de alunos matriculados em cada ciclo de estudos, em idade normal de frequência desse ciclo, entre a população residente dos mesmos escalões etários (taxa real de escolarização), aumentou, sobretudo a partir do final da década de 60, em todos os níveis (figura 4.2.1). Com exceção da Educação Pré-escolar, de frequência facultativa, a taxa de escolarização diminui à medida que avançamos no grau de escolaridade: a taxa de escolarização atinge os 100% no Ensino Primário, desde a década de 80, e, em 2010, situou-se próximo dos 94% para o 2.º Ciclo do Ensino Básico, dos 90% para o 3.º Ciclo e ultrapassou os 70% no Ensino Secundário. Devemos, no entanto, ter em conta que este indicador não traduz o efeito das retenções e desistências sobre a escolarização real, pois refere-se ao número de alunos matriculados e não àqueles que concluíram, de facto, o nível de ensino em questão. Por outro lado, e especialmente nos níveis de instrução mais elevados, existem alunos cuja idade se situa fora do intervalo considerado como normal para a frequência desse ciclo e, como tal, esses alunos não são contabilizados para determinar a escolarização nesse grau.

Mesmo entre os indivíduos mais velhos, parece existir uma tendência para o aumento da escolarização, mesmo que numa perspetiva informal: a proporção de pessoas entre os 25 e os 64 anos que participam em atividades de educação e formação¹² tem aumentado nos últimos anos e, em 2009, foi de 6,5% (figura 4.2.2). Apesar de este indicador se manter ainda em valores muito baixos, chama-nos a atenção

¹²São consideradas “todas as atividades de aprendizagem intencional ou não, desenvolvidas ao longo da vida, em contextos formais, não-formais ou informais, com o objetivo de adquirir, desenvolver ou melhorar conhecimentos, aptidões e competências, no quadro de uma perspetiva pessoal, cívica, social e/ou profissional”. (INE)

Superior, os valores mais elevados correspondem aos anos de 2006-2007 e 2007-2008, quando este indicador ultrapassou os 80.000 alunos. Nos anos seguintes, voltou a descer ligeiramente, tendo atingido os 78.609 alunos diplomados no ano de 2010 (figura 4.2.3).

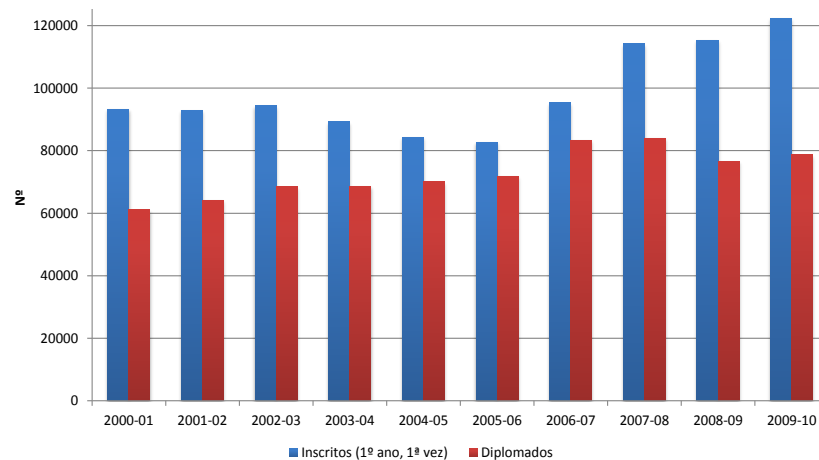


FIGURA 4.2.3. Número de inscritos e diplomados no Ensino Superior, 2000/2001 a 2009/2010

Fonte: Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais (MCTES)

4.2.2. Estrutura populacional por sexo, grupo etário e nível de escolaridade mais elevado completo. Com base nos dados obtidos nos XII, XIII e XIV Recenseamentos Gerais da População, correspondentes aos anos de 1981, 1991 e 2001, respetivamente, podemos caracterizar a evolução da estrutura da população considerando as três dimensões em estudo: sexo, idade e nível de escolaridade. Relativamente ao sexo, verifica-se que a percentagem de mulheres aumentou de aproximadamente 48% em 1981 e 1991 para 51% em 2001. O facto de a proporção de mulheres ter ultrapassado a de homens estará relacionado com o envelhecimento da população, dadas as diferenças existentes nas relações de masculinidade entre os grupos etários: nas faixas etárias mais jovens, o número de homens excede o de mulheres, como reflexo do maior número de nascimentos do sexo masculino; à medida que os indivíduos envelhecem, devido às diferenças em termos de mortalidade verificadas entre os sexos, discutidas anteriormente, o número de mulheres passa a ser superior ao de homens no mesmo escalão etário. Como tal, à medida que a proporção de idosos aumenta, cresce também a representatividade das mulheres na população.

As pirâmides apresentadas nas figuras 4.2.4, 4.2.5 e 4.2.6, permitem verificar que, como consequência da evolução das componentes demográficas, o envelhecimento da população tem ocorrido por duas vias: por um lado, com o aumento da longevidade, tem aumentado o número de pessoas acima dos 65 anos e, da mesma forma, dos idosos mais velhos (com idades superiores a 75 anos); por outro lado, a queda dos níveis de fecundidade tem conduzido à redução do número de efetivos nas idades mais jovens. Este processo tem sido agravado pela diminuição do impacto dos movimentos migratórios para o rejuvenescimento da população.

4. A EDUCAÇÃO EM PORTUGAL

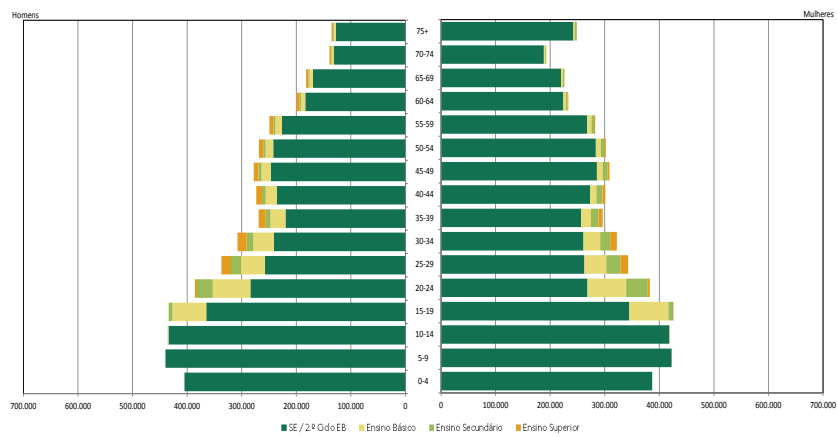


FIGURA 4.2.4. *População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 1981*

Fonte: INE, XII Recenseamento Geral da População

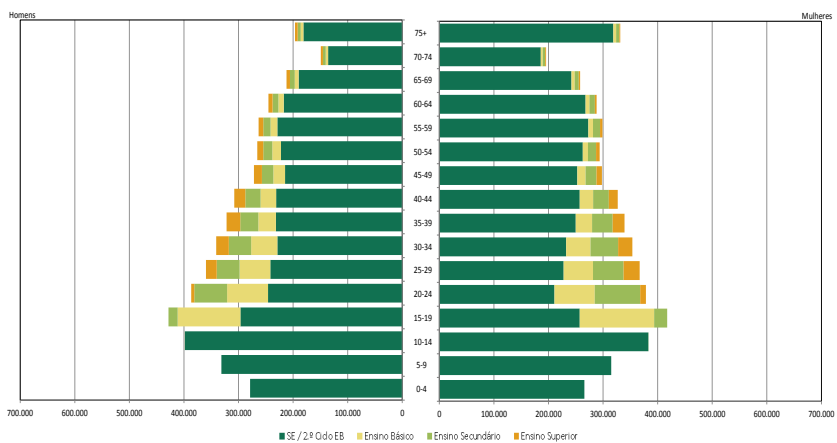


FIGURA 4.2.5. *População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 1991*

Fonte: INE, XIII Recenseamento Geral da População

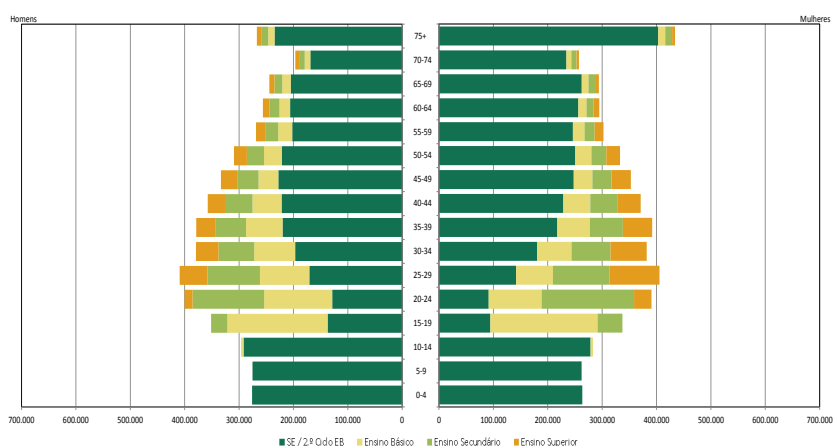


FIGURA 4.2.6. *População portuguesa por grupo etário, sexo e nível de escolaridade, Censos 2001*

Fonte: INE, XIV Recenseamento Geral da População

O aumento da escolarização é também evidente em todos os grupos etários e para ambos os sexos: em cada um dos grupos determinados por estas duas variáveis, existe uma proporção cada vez maior de indivíduos que concluíram pelo menos o Ensino Básico, tendo sido sobretudo no decorrer da década de 1991 a 2001 que se verificaram as maiores alterações. Em 1981, 89,7% dos residentes em Portugal não tinham concluído o Ensino Básico e, até 2001, esta percentagem desceu até aos 68%. A conclusão de um curso de Ensino Superior passou a ser mais frequente e, em 2001, 6,5% dos indivíduos tinham atingido este patamar máximo, quando em 1991 este valor se situava somente nos 2,9%. Nas coortes mais jovens, a escolarização parece ser distinta entre os sexos, com uma maior proporção de mulheres a concluir níveis de ensino mais elevados, relativamente à percentagem verificada entre os homens. Dado que o nível de escolarização de um indivíduo só pode manter-se constante ou aumentar, esperamos que, à medida que o tempo avança, os grupos etários mais velhos fiquem cada vez mais escolarizados, mesmo que a conclusão dos níveis de ensino ocorra somente nas idades mais jovens (neste sentido, devemos ter ainda em conta as diferenças nos padrões de mortalidade dos indivíduos por nível de escolaridade, que podem explicar variação da representatividade dos vários níveis de ensino nos diferentes grupos etários).

4.3. Objetivos e metas para a Educação

Os grandes objetivos para o desenvolvimento da escolarização da população portuguesa que se encontram atualmente em vigor, com relação ao horizonte temporal até 2020, foram estabelecidos ao abrigo do quadro estratégico para a cooperação europeia no campo da educação e formação, definido em Maio de 2009 pelo Conselho da União Europeia (*EF 2020*). Este quadro traduz-se no programa de trabalho *Educação e Formação para 2020*, baseado no programa antecessor *Educação e Formação para 2010*, e no qual se preveem quatro objetivos estratégicos comuns para os Estados-Membros: “1. tornar a aprendizagem ao longo da vida e a mobilidade uma realidade; 2. melhorar a qualidade e a eficácia da educação e da formação; 3. promover a igualdade, a coesão social e a cidadania ativa; 4. incentivar a criatividade e a inovação, incluindo o espírito empreendedor, a todos os níveis da educação e da formação”. O mesmo programa estabelece um “conjunto de níveis de referência do desempenho médio europeu” que não devem ser considerados como metas concretas a atingir por cada país até 2020, mas antes como valores de referência que permitam, a cada Estado-Membro, “ponderar, com base nas prioridades nacionais e tendo simultaneamente em conta as alterações da situação económica, como e em que medida podem contribuir para o cumprimento colectivo dos critérios de referência europeus através de medidas nacionais”. Como exemplo, entre estes níveis de referência refere-se que “até 2020, uma média de pelo menos 15% de adultos deverá participar na aprendizagem ao longo da vida”, “a percentagem de adultos de 30-34 anos com nível de ensino superior deverá ser de pelo menos 40%” ou “a percentagem de alunos que abandonam o ensino e a formação deverá ser inferior a 10%”.

No plano internacional, Portugal participou ainda, conjuntamente com os restantes países que integram a Organização de Estados Ibero-americanos, na elaboração do documento *Metas Educativas 2021: A*

Educação que queremos para a geração dos Bicentenários, que inclui um conjunto de objetivos, metas e instrumentos de avaliação regional para o desenvolvimento educativo. Entre as metas estabelecidas no âmbito deste programa, pretende-se assegurar que, até 2021, entre 60% a 90% de jovens terminam o Ensino Secundário (*Programa Educação 2015*).

Metodologia de projeção

Considerando os temas abordados nos capítulos anteriores, estamos agora em condições de estabelecer a maioria dos pressupostos e tomar as decisões necessárias à elaboração das projeções pretendidas. Seguindo o esquema proposto por Willekens (1990), devemos, em primeiro lugar, identificar o sistema em estudo: estamos interessados em elaborar projeções sobre a população residente em Portugal por sexo, grupo etário e nível de escolaridade, em intervalos quinquenais para o período 2001-2031¹³.

A análise conduzida no capítulo 3 e na secção 4.2 permitiu verificar de que forma o comportamento recente desta população se tem alterado, no que se refere às variáveis demográficas e aos indicadores de escolarização, e que modificações na estrutura populacional têm daí resultado. Recapitulando os principais pontos abordados, vimos, no que diz respeito às componentes demográficas, que: (a) os níveis de mortalidade têm vindo a diminuir em todos os grupos etários, com particular destaque para a mortalidade ocorrida durante o primeiro ano de vida, conduzindo ao aumento da esperança média de vida em todas as idades, indicador que evolui com a manutenção de um diferencial entre os sexos (sendo que as mulheres apresentam valores mais elevados de esperança de vida, relativamente aos homens); (b) os níveis de mortalidade infantil distinguem-se entre diferentes níveis de instrução da mãe, com a diminuição da mortalidade à medida que o grau de instrução da mãe aumenta, sobretudo a partir da conclusão do Ensino Secundário; (c) as mulheres portuguesas têm cada vez menos filhos e iniciam a maternidade cada vez mais tarde, estando este adiamento associado à conclusão de níveis de escolaridade superiores; (d) são as mães com níveis de escolaridade intermédios, em particular com o Ensino Básico, que têm, em média, menos filhos; (e) a contribuição das migrações para o crescimento populacional tem vindo a reduzir-se e o perfil dos migrantes no que toca ao seu grau de escolarização não parece ser totalmente homogéneo, com uma maior representatividade de pessoas com o Ensino Básico e Secundário, tanto entre os emigrantes como entre os imigrantes. Como resultado de todas estas alterações, a taxa de crescimento efetivo da população tem vindo a diminuir, aproximando-se atualmente de zero, e a população encontra-se cada vez mais envelhecida. Relativamente aos indicadores específicos sobre a educação, tem-se observado um aumento da taxa de escolarização correspondente a todos os níveis de ensino, sendo evidente o aumento da proporção de indivíduos que

¹³A intenção de iniciar o período de projeção com base em dados censitários obriga-nos a recuar até ao ano de 2001, dado que no momento de elaboração deste trabalho não se encontram ainda disponíveis os dados relativos ao XV Recenseamento Geral da População (Censos 2011) com o nível de desagregação necessário (por sexo, grupo etário e nível de escolaridade). Esta decisão apresenta a vantagem de permitir uma avaliação dos resultados das projeções produzidas por comparação com as estimativas já disponíveis para os anos 2001-2011, um intervalo de tempo bastante amplo que se reduziria caso o ano de início das projeções fosse posterior. Na secção 7.1 apresenta-se uma avaliação baseada nos dados mais recentes disponíveis, relativos ao ano de 2011. A decisão referente ao final do período de projeção prende-se com os objetivos do projeto em que este trabalho se enquadra.

completam níveis de ensino superiores em ambos os sexos e em todos os grupos etários que compõem a população.

De que forma poderemos, então, obter estimativas sobre a evolução futura da população portuguesa, considerando o seu grau de instrução? Antes de mais, interessa indicar que dados se encontram disponíveis para a elaboração das projeções: relativamente à população residente em Portugal, por sexo, grupo etário e qualificação académica, dispomos dos dados censitários referentes aos anos de 1981, 1991 e 2001, e de estimativas anuais obtidas através do Inquérito ao Emprego, elaborado pelo INE, I.P., referentes à população com mais de 15 anos (por grupos etários quinquenais) para os anos entre 1998 e 2010; quanto aos eventos relevantes ocorridos, dispomos do número anual de nados-vivos por grupo etário e nível de escolaridade mais elevado completo da mãe, do número de óbitos com menos de 1 ano por sexo e nível de escolaridade mais elevado completo da mãe, e do saldo migratório total.

Com base na análise efetuada nos capítulos anteriores, sabemos que os modelos de projeção multiestado são os mais adequados para considerar populações divididas segundo diferentes características e as interações entre estes subgrupos. Pretendemos assegurar que as projeções englobam uma perspetiva dinâmica da evolução demográfica, aproximando-se o mais possível do que ocorre na realidade, e assumiremos, como tal, que tanto os níveis de fecundidade, mortalidade e migrações, bem como de transições entre categorias de escolaridade podem sofrer modificações ao longo do tempo. A abordagem ideal, neste caso, seria adotar um modelo dinâmico multi-idade e multiestado (ver secção 1.4.4) para modelar e projetar a população; no entanto, os substanciais requisitos de dados associados à utilização destes modelos limitam esta opção. Optar-se-á, em vez disso, por uma abordagem sugerida por Schoen (2006b), e cuja aplicação se adequa mais facilmente aos dados de que dispomos. Recorreremos a modelos matriciais discretos, com base na metodologia de componentes por coorte adaptada a uma população multidimensional, assumindo que as taxas que comandam a dinâmica demográfica se mantêm constantes em cada quinquénio de projeção. A estrutura e características da população de partida e as taxas relativas a cada quinquénio são, desta forma, os fatores que determinam os resultados das projeções.

Dada a inexistência de uma metodologia padrão para realizar projeções probabilísticas sobre populações multidimensionais e o uso generalizado das projeções determinísticas, particularmente entre os organismos e investigadores que produzem projeções por nível de escolaridade, opta-se pelo recurso a estas, através da elaboração de cenários alternativos sobre a evolução futura das várias componentes.

5.1. População de partida

A população na qual se baseiam as projeções foi obtida do seguinte modo: em primeiro lugar, obtiveram-se as estimativas do INE, I.P., referentes à população residente em Portugal em 31 de dezembro de 2000 (Estimativas Anuais da População Residente), apresentada na tabela A.0.1 do

apêndice A, estimada com base na população recenseada em 12 de março de 2001 (Censos 2001) e recuada ao final do ano anterior¹⁴ (Carrilho, 2005); posteriormente, foi aplicada a cada grupo desta população (por sexo e grupo etário) a estrutura verificada nos Censos 2001 relativa à qualificação académica (nível de instrução mais elevado completo), apresentada na tabela A.0.2. Pretendemos que o instante de referência das projeções seja o dia 1 de janeiro de 2001 e, como tal, assumimos que a população residente nesse dia é igual à estimada pelo INE, relativa a 31 de dezembro de 2000 e, ainda, que não existiram alterações na estrutura por qualificação académica entre o momento censitário e o dia de referência¹⁵. Ficamos, assim, com a população de partida para as projeções, dividida por sexos, grupos etários quinquenais (entre os 0-4 anos e os 75 e mais anos, representados por $z = 1, \dots, 16$, respetivamente) e nível de escolaridade mais elevado completo (ver tabela A.0.3).

5.2. Desenho do modelo de projeção

5.2.1. Especificação do espaço dos estados. Considerando a estrutura do Sistema Educativo Português e as reformas que têm vindo a ser implementadas, optou-se pela divisão dos indivíduos de cada sexo e grupo etário em quatro categorias, consoante o seu nível de escolaridade mais elevado completo:

- (1) *Sem escolaridade/2.º Ciclo Ensino Básico*, grupo que abrange os indivíduos que não completaram nenhum grau de ensino formal e aqueles que completaram somente o 1.º Ciclo do Ensino Básico, anteriormente designado por Ensino Primário, ou o 2.º Ciclo do Ensino Básico¹⁶;
- (2) *Ensino Básico*, englobando os indivíduos que completaram o atual 3.º Ciclo do Ensino Básico (9.º ano de escolaridade), antigo Ensino Básico Unificado;
- (3) *Ensino Secundário*, para aqueles que completaram o 12.º ano de escolaridade, antigamente designado como Ensino Propedêutico, ou Cursos Médios (na categorização utilizada nos Censos 1981, abrangemos os Cursos de Índole Profissional e Artísticos e os Cursos Médios, Enfermagem e Profissionais);
- (4) *Ensino Superior*, categoria que abrange os indivíduos que completaram algum Curso Superior, atingindo, pelo menos, o grau de Bacharel ou Licenciado.

¹⁴ Este ajuste foi feito pelo INE, com base nos valores de natalidade e mortalidade verificados entre os dois instantes.

¹⁵ Dado que o momento em que ocorre, habitualmente, a conclusão do ano letivo e a consequente transição de indivíduos entre níveis de escolaridade não se situa entre 1 de janeiro e 12 de março, este pressuposto é consistente com a realidade, desde que assumamos também a ausência de diferenciais na mortalidade e migrações ocorridas entre os dois momentos.

¹⁶ Nos grupos etários dos 0-4 e 5-9 anos, o primeiro estado corresponde somente aos casos em que não existe nenhum grau de escolaridade ou, numa reduzida proporção das crianças com 5-9 anos, a indivíduos com o 1.º Ciclo do Ensino Básico: de acordo com a estrutura do sistema educativo português, descrita no capítulo anterior, não é possível completar o 2.º Ciclo do Ensino Básico antes dos 10 anos.

A decisão de agrupar os indivíduos sem escolaridade formal e aqueles que completaram até o 2.º Ciclo do Ensino Básico na mesma categoria prende-se com dois fatores principais:

- (a) Quanto mais elevado for o número de grupos de escolaridade considerados, mais complexo será o modelo de projeção e os requisitos iniciais de dados;
- (b) O regime de escolaridade obrigatória para as crianças e jovens entre os 6 e os 15 anos foi estabelecido pelo Decreto-Lei n.º 301/93, de 31 de Agosto, e virá a ser alargado até aos 18 anos, tendo sido já publicada a lei que assim o determina; neste sentido, parece-nos que a projeção do número de indivíduos que apenas concluíram o 1.º ou 2.º Ciclos do Ensino Básico terá pouco interesse, já que, à partida, nos últimos anos e, sobretudo, no horizonte temporal considerado para a projeção, todos os indivíduos deverão completar esses níveis de ensino.

A correspondência entre as categorias utilizadas e a Classificação Internacional Normalizada da Educação (*International Standard Classification of Education* ou ISCED), desenvolvida pela UNESCO (2006) e destinada a permitir a comparação de estatísticas e de políticas educativas entre sistemas educativos diferentes, é apresentada na tabela seguinte:

TABELA 5.2.1. *Correspondência entre categorias de escolaridade e níveis ISCED (1997)*

Categoria	Nível ISCED
Sem escolaridade/2.º Ciclo Ensino Básico	1
Ensino Básico	2
Ensino Secundário	3
Ensino Superior	4, 5 e 6

5.2.2. Especificação dos eventos. Dada a definição anterior do espaço dos estados, e considerando um sistema em que é possível ocorrerem entradas por via do nascimento ou imigração e saídas por morte ou emigração, podemos estabelecer a seguinte matriz de eventos em cada intervalo de cinco anos:

TABELA 5.2.2. *Matriz de eventos para o modelo de projeção por escolaridade*

Legenda: EI - Evento interno; EE - Evento externo; O - Não evento (indivíduo fica na mesma categoria); * - Evento não relevante; Célula vazia - Evento impossível.

De:	Para:				
	SE/2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	Morte/Emigração
SE/2.º Ciclo EB	O	EI	EI		EE
Ensino Básico		O	EI		EE
Ensino Secundário			O	EI	EE
Ensino Superior				O	EE
Nascimento	EE				*
Imigração	EE	EE	EE	EE	*

Esta matriz foi construída com base nos seguintes pressupostos:

- (i) As transições de um nível de escolaridade para outro podem ocorrer apenas numa única direção, passo a passo, e seguindo uma sequência pré-definida; em resultado, um indivíduo que tenha concluído um curso de Ensino Superior terá de ter passado por todos os níveis anteriores e irá manter-se neste nível de escolaridade no futuro, mesmo que as suas capacidades ou competências reais se alterem;
- (ii) Como se pretende projetar a população em intervalos de cinco anos, podemos admitir que nesse intervalo de tempo existe a possibilidade de um indivíduo completar o Ensino Básico e o Ensino Secundário, transitando do primeiro para o terceiro estado;
- (iii) Todos os recém-nascidos ocupam, obviamente, o primeiro estado, independentemente do grupo etário da mãe ou do estado que esta ocupa no momento do nascimento;
- (iv) Considera-se que a saída do sistema (por morte ou emigração) para os indivíduos que acabaram de entrar (por nascimento ou imigração) é um evento sem relevância, já que estes indivíduos não chegam a entrar, de facto, no sistema.

5.2.3. Projeção das componentes demográficas e da escolaridade. Um dos principais desafios na elaboração de projeções demográficas é a projeção de valores para as taxas de fecundidade, mortalidade, para as migrações e, em particular no caso das populações multidimensionais, para as transições entre estados. Nos modelos multiestado deve considerar-se a existência de diferenciais de fecundidade, mortalidade e migrações entre os vários estados, aspeto que contribui para aproximar o modelo à dinâmica real da população e pode, por isso, determinar uma maior qualidade dos resultados, em comparação com os modelos simples.

5.2.3.1. Mortalidade. A medição direta da mortalidade por nível de escolaridade exige um sistema de registo de óbitos completo e fiável, juntamente com informação sobre a escolaridade dos falecidos e as correspondentes populações em risco; contudo, tais dados empíricos estão limitados a apenas alguns países industrializados e não se encontram disponíveis em Portugal, dado que as certidões de óbito não incluem o registo do nível de instrução do falecido. Como tal, dificulta-se o estudo e projeção da mortalidade considerando, de forma diferenciada, os subgrupos definidos pelo grau de instrução. À semelhança de Goujon et al. (2007), recorreremos a informação relativa à taxa de mortalidade infantil por grupo de escolaridade da mãe¹⁷ (indicadores já calculados e apresentados na figura 3.1.4 e tabela A.0.4) para introduzir diferenciais na mortalidade por nível de escolaridade, nos restantes grupos etários.

¹⁷Importa considerar que a taxa de mortalidade infantil para um determinado ano t , calculada através da expressão $TMI_t = \frac{O_0[t]}{B[t]}$, onde $O_0[t]$ é o número de óbitos com idade inferior a 1 ano, no ano t , e $B[t]$ é o total de nascimentos no mesmo período, não é uma taxa no sentido em que nos temos referido, pois tem contagens no denominador, e também não é uma probabilidade, porque o numerador depende de nascimentos ocorridos tanto no ano t como no ano $t - 1$ (ou seja, os óbitos no numerador não correspondem, na sua totalidade, aos elementos considerados no denominador) (Carvalho, 2010). Ainda assim, esta é uma aproximação frequentemente utilizada e é o indicador possível de calcular com os dados disponíveis.

Partindo das tábuas abreviadas de mortalidade disponibilizadas pelo INE, I.P., é sobre os quocientes de mortalidade entre as idades exatas x e $x + n$, representados por ${}_nq_x$ (ver tabela A.0.5), que são estabelecidos os diferenciais de mortalidade por nível de instrução. Estes valores são fornecidos para cada sexo e em relação a períodos bienais, entre 1991-2001 e 2008-2010. Assim, para o primeiro quinquénio de projeção (2001-2006) foi calculado o quociente de mortalidade médio em cada intervalo de idades para os biénios 2000-2002 a 2005-2007. As relações obtidas entre as taxas de mortalidade infantil calculadas para cada grupo de escolaridade da mãe e a taxa de mortalidade infantil para o conjunto da população foram depois utilizadas para estimar os quocientes de mortalidade específicos para o sexo, idade e nível de escolaridade. A partir destes quocientes estimam-se, para cada sexo e nível de instrução, os seguintes valores, de acordo com as expressões apresentadas¹⁸:

TABELA 5.2.3. *Funções de mortalidade estimadas para cada sexo e nível de escolaridade*

Função	Designação	Expressão
${}_np_x$	Probabilidade de sobrevivência entre as idades exatas x e $x + n$	${}_np_x = 1 - {}_nq_x$
ℓ_x	Sobreviventes à idade exata x	$\ell_0 = 100000$ e $\ell_{x+n} = \ell_x \cdot {}_np_x$
${}_nd_x$	Óbitos entre as idades exatas x e $x + n$	${}_nd_x = \ell_x \cdot {}_nq_x$
${}_nL_x$	Pessoa-anos vividos entre as idades exatas x e $x + n$	${}_nL_x = n \cdot \ell_{x+n} + \frac{n}{2} \cdot {}_nd_x$
T_x	Pessoa-anos vividos após a idade x	$T_x = \sum_{a=x}^{\infty} {}_nL_x$

Obtendo as estimativas correspondentes, podemos calcular a proporção de sobrevivência para cada grupo etário, ou seja, a proporção de pessoas no i -ésimo grupo etário no momento inicial t que sobrevivem de modo a pertencer ao grupo etário $i + 1$ no momento $t + n$, específica por sexo e nível de escolaridade. Estas proporções são estimadas a partir das seguintes expressões:

¹⁸ Assume-se a ausência de transições entre níveis de escolaridade; admite-se também que as mortes no intervalo $[x, x+n[$ se dão a meio do intervalo e, como tal, ${}_na_x = \frac{n}{2}$.

TABELA 5.2.4. Expressões de cálculo da proporção de sobrevivência, para cada sexo e nível de escolaridade, por grupo etário

Grupo etário (i)	Idades $[x, x + n[$	Expressão	Significado
0	$[0, 1[$	$s_0 = \frac{{}_iL_0 + {}_4L_1}{5 \cdot \ell_0}$	Proporção de indivíduos que tem 0-4 anos no instante $t + n$, de entre todos os nascimentos ocorridos no quinquénio $[t, t + n[$
1	$[1, 5[$	$s_1 = \frac{{}_5L_5}{({}_1L_0 + {}_4L_1)}$	Proporção de indivíduos no grupo etário 0-4 anos em t que sobrevivem e pertencem ao grupo 5-9 anos em $t + n$
2 a 14	$[5, 10[$ a $[65, 70[$	$s_i = \frac{{}_nL_{x+n}}{{}_nL_x}$	Proporção de pessoas no i -ésimo grupo etário no momento inicial t que sobrevivem e pertencem ao grupo etário $i + 1$ no momento $t + n$
15	$[70, 75[$	$s_{15} = \frac{T_{75}}{T_{70}}$	Proporção de indivíduos no grupo etário 70-74 anos em t que sobrevivem e pertencem ao grupo 75+ anos em $t + n$

Os resultados são apresentados na tabela 5.3.3. Seguindo estes passos, não obtemos uma verdadeira tabela de mortalidade multiestado (ver secção 1.4.2), pois não são consideradas, para já, as eventuais transições entre níveis de escolaridade; obtemos, em vez disso, quatro tabelas de mortalidade distintas para cada sexo, uma para cada categoria de escolaridade e refletindo já os diferenciais de mortalidade por grau de instrução, sendo que as proporções de sobrevivência calculadas correspondem não só à sobrevivência para o grupo etário seguinte, mas também à permanência no mesmo estado.

5.2.3.2. *Fecundidade.* Para cada nível de escolaridade i e grupo etário da mãe j ¹⁹, a Taxa de Fecundidade específica foi calculada, para cada um dos anos t entre 2001 e 2006, como:

$$\tilde{b}_{ijt} = \frac{\text{nº de nados-vivos em mulheres com nível de escolaridade } i \text{ e grupo etário } j, \text{ no ano } t}{\text{pessoa-anos vividos por mulheres com nível de escolaridade } i \text{ e grupo etário } j, \text{ no ano } t}$$

onde o denominador é aproximado pela população média residente com as mesmas características²⁰. Os dados relativos à população feminina e aos nados vidos estão apresentados nas tabelas A.0.6 e A.0.7, respetivamente. O total das taxas obtidas para os vários anos foi utilizado para estimar o número de nascimentos verificados no quinquénio.

A média das relações de masculinidade no primeiro grupo etário (0-4 anos) verificadas entre os anos de 2001 e 2006 (ver tabela A.0.8) foi utilizada para distribuir o total de nascimentos estimados para

¹⁹ Foram consideradas somente as mulheres em idade reprodutiva, admitindo que este intervalo se estende entre os 15 e os 49 anos.

²⁰ A semelhança dos pressupostos assumidos no tratamento da mortalidade, admite-se que os nascimentos se encontram igualmente distribuídos ao longo do quinquénio (ou, equivalentemente, ocorrem exatamente a meio do intervalo).

cada intervalo pelos dois sexos. Foi ainda considerada a possibilidade de morte no primeiro ano de vida, com base na proporção de sobrevivência específica por sexo, obtida para os indivíduos sem escolaridade no grupo etário 0, através da expressão apresentada na secção anterior (tabela 5.2.4).

5.2.3.3. *Migrações.* A análise e projeção da componente migratória é sujeita à escassez de dados detalhados sobre a população migrante, nomeadamente no que se refere à sua estrutura por nível de escolaridade, bem como à dificuldade de previsão do comportamento desta variável, inerente à possibilidade de poder ocorrer em diferentes sentidos (emigração ou imigração) e por ser fortemente afetada pela conjuntura nacional e internacional, que podem evoluir de forma inesperada. Como referido por Alho e Keilman (2010), as migrações são habitualmente descritas em termos de valores absolutos, dada a dificuldade associada à definição da população em risco de imigração, e não em relação a taxas. Nesse caso, é a diferença entre as imigrações e as emigrações num dado intervalo que determina o efeito da migração externa na evolução futura da população e, não existindo a necessidade de registar o número de imigrantes e emigrantes separadamente, pode recorrer-se somente ao saldo migratório.

O INE, I.P., disponibiliza estimativas anuais do saldo migratório verificado em Portugal, englobando os indivíduos de ambos os sexos, de todos os grupos etários e níveis de instrução (ver tabela A.0.9). A distribuição deste valor por sexo e grupo etário foi conduzida com base na estrutura dos movimentos migratórios internacionais para Portugal, apresentada pela Organização das Nações Unidas (UN, 2011). A distribuição subsequente por nível de escolaridade partiu do pressuposto de que a estrutura verificada nos migrantes não difere substancialmente da verificada na população residente²¹. Como tal, a proporção de migrantes em cada nível de escolaridade foi estabelecida como igual à dos residentes no mesmo ano (para o grupo etário dos 10-14 anos foi sempre utilizada a estrutura verificada nos Censos 2001, sendo este o único ano em que dispomos de dados para este grupo etário).

5.2.3.4. *Transições entre níveis de escolaridade.* A conclusão de um nível de escolaridade e a inerente transição para um estado distinto nunca é um evento certo, e estas transições não se relacionam diretamente com o tempo, ao contrário do que acontece com a progressão da idade. Como tal, a modelação multiestado, centrada na estimação das intensidades ou das probabilidades de transição entre estados, exige a obtenção de dados específicos, relativos não só ao número de transições, mas também ao momento em que ocorrem. No nosso caso concreto, podemos focar-nos na estimação das probabilidades de transição relativas a cada um dos períodos de 5 anos em que se divide o intervalo de projeção. Antes de mais, é necessário ter em conta que o cálculo de tais probabilidades deve ter por base a distribuição da população pelos vários níveis de escolaridade, por sexo e grupo etário, em pelo menos dois pontos no tempo. Se dispusermos somente de uma observação, num único ponto no

²¹ Os gráficos 3.1.9 e 3.1.10, apresentados no capítulo 3, permitem fundamentar este pressuposto; o teste desta diferença, em termos estatísticos, é condicionado pela forma como é feito o agrupamento dos residentes, dos emigrantes e dos imigrantes, em relação ao seu nível de escolaridade, que é distinta nos três casos.

tempo, torna-se impossível distinguir entre o padrão etário de aumento da escolarização e as possíveis diferenças na progressão entre coortes (Goujon, 2008). Dado que a escolarização tem vindo, de forma geral, a aumentar ao longo do tempo e, como tal, as coortes mais velhas tendem a ser menos escolarizadas relativamente às coortes mais jovens, a distribuição transversal por sexo, grupo etário e nível de escolaridade, indica, habitualmente, que a escolarização aumenta inicialmente com a idade e diminui depois nos indivíduos de idade mais avançada. Estas diferenças inter-coorte são ajustadas quando observamos a população em diferentes pontos no tempo.

Considerando as idades esperadas de transição, de acordo com a estrutura do Sistema Educativo Português, discutida anteriormente, admite-se que as transições do estado Sem escolaridade/2.º Ciclo EB para o Ensino Básico ocorrem em torno dos 15 anos de idade, sendo por isso calculadas taxas de transição entre estes níveis para os grupos etários 10-14 e 15-19 anos. Já as transições do Ensino Básico para o Ensino Secundário (que acontecem, teoricamente, por volta dos 18 anos) poderão ocorrer nos grupos etários 15-19 e 20-24 anos e, por fim, assume-se que as transições do Ensino Secundário para o Ensino Superior acontecem nos grupos 20-24 e 25-29 anos. Ainda com base na organização do Sistema de Ensino e devido ao facto de os indivíduos não poderem retroceder a um nível de escolaridade mais baixo, admite-se que não ocorrem transições entre os níveis de escolaridade após os 30 anos de idade²².

Em relação à informação que esteja disponível para a estimação, interessa determinar, por exemplo: se se refere a unidades estatísticas individuais ou a grupos de indivíduos; se foi recolhida em pontos discretos no tempo ou se houve um registo contínuo; se segue um esquema de coorte, período ou período-coorte; se se baseia em observações para toda a população ou somente para uma amostra ou se os dados foram recolhidos prospetiva ou retrospectivamente, já que o plano observacional vai influenciar a decisão sobre a metodologia mais adequada para a estimação (Gill e Keilman, 1990). Se for observada a trajetória de vida de um grupo de n indivíduos durante um determinado intervalo de tempo, ou seja, se os instantes exatos de transição forem conhecidos para cada um dos elementos da amostra no decorrer desse intervalo (o que alguns autores designam como micro-dados), então a estimação das taxas de transição entre dois instantes pode basear-se no estimador de máxima verosimilhança, válido sob o pressuposto de um processo homogéneo:

$$\tilde{m}_{ij}(t, t+h) = \frac{\tilde{d}_{ij}(t, t+h)}{\tilde{L}_i(t, t+h)} = \frac{\sum_{l=1}^n N_{ij}^l(t, t+h)}{\sum_{l=1}^n T_i^l(t, t+h)} \quad i \neq j$$

onde $N_{ij}^l(t, t+h)$ indica o número total de transições do estado i para o estado j feitas pelo indivíduo l no intervalo $[t, t+h[$, e T_i^l é o tempo total passado por l no estado i , no mesmo intervalo (Mahlmann, 2006).

²²Como apresentado no gráfico 4.2.2, a percentagem de pessoas entre os 25 e os 64 anos que participam em atividades de educação e formação tem vindo a aumentar; contudo, esta proporção é ainda muito baixa e inclui a participação em atividades de aprendizagem “em contextos formais, não-formais ou informais”, podendo conduzir, ou não, a transições entre os estados considerados. Assim, opta-se por assumir que a percentagem de pessoas que transitam entre os grupos considerados após os 30 anos de idade é residual e pode ser ignorada.

Por outro lado, este plano observacional permite também estimar facilmente as probabilidades de transição, admitindo novamente que a intensidade de transição se mantém constante ao longo do tempo (probabilidades de transição estacionárias): a probabilidade de transitar a partir de qualquer estado i para o estado j é igual à proporção de indivíduos que estavam no estado i num dado instante t e passaram a estar no estado j no instante $t + h$, de entre todos os indivíduos que estavam no estado i em t :

$$\tilde{p}_{ij}(t, t + h) = \frac{\tilde{\ell}_{ij}(t + h)}{\tilde{\ell}_i(t)} = \frac{\sum_{l=1}^n Y_{ij}^l(t, t + h)}{\sum_{l=1}^n Y_i^l(t)}$$

onde $Y_{ij}^l(t, t + h)$ adquire o valor 1 se o indivíduo l está no estado i no instante t e no estado j no instante $t + h$, e $Y_i^l(t)$ tem o valor 1 se o indivíduo l está no estado i no instante t . Estes cálculos podem ser efetuados em relação a subgrupos específicos da população, nomeadamente para cada coorte determinada pelo sexo e grupo etário (nesse caso, a amplitude de cada grupo etário deverá coincidir com a de cada intervalo de tempo, h , avançando os indivíduos para o grupo etário seguinte quando atingimos o final do intervalo). Anderson e Goodman (1957) mostram que este é um estimador de máxima verosimilhança que é consistente mas enviesado, com o viés a tender para zero à medida que a dimensão da amostra aumenta. Assim, é possível estimar uma matriz de transição consistente recorrendo a uma amostra suficientemente grande (Jones, 2005). Em conclusão, quando os indivíduos são observados separadamente, seja continuamente ou em pontos discretos no tempo, é possível obter diretamente a equação de verosimilhança e desenvolver procedimentos de inferência eficientes.

Neste trabalho, no entanto, e à semelhança do que acontece em muitas aplicações práticas, encontram-se disponíveis apenas os chamados dados agregados, em que não são observadas as transições individuais mas somente o número de indivíduos que ocupa cada estado em diferentes instantes de observação, o que conduz à necessidade de recorrer a métodos de estimação mais complexos.

Diversos autores têm apresentado metodologias destinadas ao tratamento de dados agregados no âmbito dos processos estocásticos e, em particular, das Cadeias de Markov. Frequentemente, são considerados processos homogêneos e de tempo discreto, com recurso a estimação por mínimos quadrados ou máxima verosimilhança. Para os casos em que dispomos do número agregado de transições (por cada coorte de indivíduos) entre cada um dos estados, para determinado período, Davis et al. (2002) propuseram um procedimento de estimação para processos de Markov não-homogêneos e de tempo discreto, por mínimos quadrados ponderados. Heathcote e Puza (2005) utilizaram uma técnica de estimação de probabilidades de estado de um processo de tempo discreto e espaço de estados finitos, discutida anteriormente por Davis et al. (2001), partindo de dados agregados transversais, em que não se conhecem quantas transições foram efetuadas. No entanto, esta abordagem não segue o pressuposto de Markov, não relaciona a ocupação de cada um dos estados com a distribuição verificada em instantes anteriores e não permite, por isso, o cálculo das probabilidades de transição, condicionadas ao estado ocupado num momento anterior. Também Moffitt (1993) apresentou metodologias para a

identificação e estimação de modelos dinâmicos a partir de um conjunto de dados transversais independentes, relativos a dois ou mais pontos no tempo. Tendo considerado a possibilidade de recorrer a um modelo de Markov, o autor abordou somente o caso em que o espaço de estados é constituído por apenas dois estados e em que se encontra disponível informação sobre um conjunto de covariáveis que são utilizadas para estimar as probabilidades de transição. Esta abordagem foi mais tarde discutida por Pelzer et al. (2001). Mais recentemente, Crowder e Stephens (2011) propuseram métodos baseados em funções geradoras de probabilidade para estimar as probabilidades de transição num processo de Markov homogêneo, com base em dados agregados (que designam no seu trabalho por macro-dados) mas sem que fossem consideradas as entradas de indivíduos no sistema. Gouno et al. (2011) sugeriram novos métodos baseados em técnicas alternativas aos modelos de Markov, que se baseiam na aplicação de um algoritmo EM e da sua versão estocástica (Monte Carlo EM – MCEM), através da estimação dos tempos de permanência em cada um dos estados (dados não observados). Não consideraram, da mesma forma, a possibilidade de entrarem elementos no sistema.

Ainda na década de 80, Kalbfleisch et al. (1983) desenvolveram procedimentos baseados em mínimos quadrados condicionais e estimação por máxima verosimilhança aproximada, para cadeias de Markov homogêneas e de tempo contínuo no caso em que se dispõe unicamente de dados agregados, admitindo a possibilidade de entrada de indivíduos no sistema. Kalbfleisch e Lawless (1984) apresentaram ainda uma forma análoga de tratar dados agregados, agora em tempo discreto, recorrendo a mínimos quadrados ponderados²³. No entanto, por pretendermos estimar probabilidades de transição num intervalo de tempo quinquenal, específicas para cada grupo etário em que as mesmas ocorrem, precisamos de comparar as proporções de indivíduos dum determinado grupo etário, em cada nível de escolaridade, no momento t , com as proporções no grupo etário seguinte, no momento $t + 5$. Desta forma, seguindo a metodologia apresentada por Kalbfleisch e Lawless (1984) ficaríamos apenas com duas observações (em t e em $t + 5$) relativas a cada proporção, o que resulta num número de observações inferior ao total de parâmetros a estimar²⁴.

Perante estas limitações, a estimação das probabilidades de transição teve de basear-se num procedimento mais simples, fazendo uso da hierarquia presente entre os níveis de escolaridade e da unidirecionalidade da passagem pelos vários estados. Seja $p_{it}(x, x + n)$ a probabilidade não condicional dos indivíduos entre as idades x e $x + n$ no instante t estarem no estado i . O vetor linha que contém a distribuição de probabilidades entre os vários estados para os indivíduos de idades entre x e $x + n$ no instante t é representado por $\mathbf{p}_t(x, x + n)$. Seja ainda $p_{ijt}(x, x + n)$ a probabilidade condicional dos indivíduos no estado i entre as idades x e $x + n$ no instante t estarem no estado j exatamente n anos depois (correspondentes às probabilidades de transição da Cadeia de Markov, apresentadas na

²³Esta metodologia é descrita no apêndice D, por ser aquela que se considera que seria a mais adequada aos tipo de dados de que dispomos, caso não pretendêssemos seguir um esquema período-coorte.

²⁴A aplicação do método descrito por Kalbfleisch e Lawless (1984) deveria basear-se em dados relativos à ocupação dos vários estados em instantes consecutivos, por exemplo ano a ano, mas relativamente a uma mesma coorte de indivíduos (por exemplo, utilizar as proporções de indivíduos em cada nível de escolaridade entre os 15 e os 19 anos em 2001, entre os 16 e os 20 anos em 2002, entre os 17 e os 21 anos em 2003 e assim sucessivamente até ao final do quinquénio, para estimar as transições referentes ao grupo etário 15-19 anos em 2001, que têm 20-24 anos em 2006).

equação 3). Então, teremos que (Willekens, 2006; Müller, 2007):

$$p_{j,t+n}(x+n, x+2n) = \sum_i p_{it}(x, x+n) p_{ijt}(x, x+n)$$

ou

$$\mathbf{p}_{t+n}(x+n, x+2n) = \mathbf{p}_t(x, x+n) \mathbf{P}_t(x, x+n)$$

em que $\mathbf{P}_t(x, x+n)$ é a matriz de probabilidades de transição $p_{ijt}(x, x+n)$.

Considerando apenas as transições relevantes neste trabalho, partindo dos pressupostos descritos nas secções anteriores, temos que:

$$\begin{aligned} p_{2,t+5}(10, 15) &= p_{1,t}(5, 10) \cdot p_{1,2,t}(5, 10) \\ p_{2,t+5}(15, 20) &= p_{1,t}(10, 15) \cdot p_{1,2,t}(10, 15) + p_{2,t}(10, 15) \cdot (1 - p_{2,3,t}(10, 15)) \\ p_{3,t+5}(15, 20) &= p_{1,t}(10, 15) \cdot p_{1,3,t}(10, 15) + p_{2,t}(10, 15) \cdot p_{2,3,t}(10, 15) \\ p_{3,t+5}(20, 25) &= p_{2,t}(15, 20) \cdot p_{2,3,t}(15, 20) + p_{3,t}(15, 20) \cdot (1 - p_{3,4,t}(15, 20)) \\ p_{4,t+5}(20, 25) &= p_{3,t}(15, 20) \cdot p_{3,4,t}(15, 20) \\ p_{4,t+5}(25, 30) &= p_{3,t}(20, 25) \cdot p_{3,4,t}(20, 25) + p_{4,t}(20, 25) \end{aligned}$$

A proporção de pessoas no estado i entre as idades x e $x+n$ no instante t , que representaremos por $\tilde{p}_{it}(x, x+n)$, é um estimador centrado de $p_{it}(x, x+n)$. Assim:

$$\begin{aligned} \tilde{p}_{1,2,t}(5, 10) &= \frac{\tilde{p}_{2,t+5}(10, 15)}{\tilde{p}_{1,t}(5, 10)} = \tilde{p}_{2,t+5}(10, 15) \\ \tilde{p}_{3,4,t}(15, 20) &= \frac{\tilde{p}_{4,t+5}(20, 25)}{\tilde{p}_{3,t}(15, 20)} \\ \tilde{p}_{3,4,t}(20, 25) &= \frac{\tilde{p}_{4,t+5}(25, 30) - \tilde{p}_{4,t}(20, 25)}{\tilde{p}_{3,t}(20, 25)} \\ \tilde{p}_{2,3,t}(15, 20) &= \frac{\tilde{p}_{3,t+5}(20, 25) - \tilde{p}_{3,t}(15, 20) + \tilde{p}_{4,t+5}(20, 25)}{\tilde{p}_{2,t}(15, 20)} \end{aligned}$$

em que $\tilde{p}_{ijt}(x, x+n)$ é o estimador de $p_{ijt}(x, x+n)$ que iremos utilizar.

Por termos um número de incógnitas superior ao número de equações, a estimação das restantes probabilidades de transição não é possível unicamente a partir das proporções de estado. Visto que as estimativas tanto de $p_{1,2,t}(10, 15)$ como de $p_{1,3,t}(10, 15)$ podem ser obtidas a partir de uma estimativa de $p_{2,3,t}(10, 15)$, opta-se por assumir que:

$$\tilde{p}_{2,3,t}(10, 15) = TRE_{E.Secundário}(t, t+5) \cdot (1 - TRD_{E.Secundário}(t, t+5))$$

onde $TRE_{E.Secundário}(t, t+5)$ representa a taxa real de escolarização média no Ensino Secundário para o período $(t, t+5)$ e $TRD_{E.Secundário}(t, t+5)$ representa a média da taxa de retenção e desistência no Ensino Secundário para o mesmo período. Esta expressão resulta dos seguintes pressupostos: a) a

conclusão de um nível de ensino exige, em primeiro lugar, a inscrição nesse ciclo de estudos, sendo a percentagem de alunos matriculados em cada ciclo de estudos, em idade normal de frequência desse ciclo, dada pela taxa real de escolarização; b) para concluir o ciclo de estudos é igualmente necessária aprovação e os alunos que não desistem ou ficam retidos são aqueles que são aprovados; c) existe independência entre a proporção de alunos matriculados num ciclo de estudos e a proporção de alunos retidos ou desistentes. Assim, $\tilde{p}_{2,3,t}(10, 15)$ representa a proporção de alunos entre os 10 e os 14 anos em t que, nos 5 anos seguintes, estão inscritos e são aprovados no Ensino Secundário, passando a pertencer ao estado correspondente a este nível, no instante $t+5$, quando se situam já no grupo etário seguinte. Os valores de taxa real de escolarização e taxa de retenção e desistência, disponibilizados pelo Gabinete de Estatística e Planeamento da Educação do Ministério da Educação, são relativos ao conjunto de ambos os sexos, o que resulta em estimativas de $p_{2,3,t}(10, 15)$ iguais para homens e mulheres.

Com base nestes pressupostos, obtêm-se facilmente os estimadores relativos às restantes probabilidades de transição:

$$\tilde{p}_{1,2,t}(10, 15) = \frac{\tilde{p}_{2,t+5}(15, 20) - \tilde{p}_{2,t}(10, 15) \cdot (1 - \tilde{p}_{2,3,t}(10, 15))}{\tilde{p}_{1,t}(10, 15)}$$

$$\tilde{p}_{1,3,t}(10, 15) = \frac{\tilde{p}_{3,t+5}(15, 20) - \tilde{p}_{2,t}(10, 15) \cdot \tilde{p}_{2,3,t}(10, 15)}{\tilde{p}_{1,t}(10, 15)}$$

Importa destacar que estas proporções de transição levam em consideração a possibilidade de saída da população por morte ou emigração, em cada intervalo de cinco anos, mas ignoram a hipótese de entrada por imigração. Como tal, a estimação de $p_{ijt}(x, x+n)$ para $i = j$ deve levar em consideração as proporções de sobrevivência apresentadas na secção 5.2.3.1, tal como se refere na secção seguinte.

5.2.4. Estimação dos efetivos populacionais. A estimação dos parâmetros relativos às componentes demográficas mortalidade, fecundidade e migrações, e das probabilidades de transição, permite, a partir da população inicial, estimar o número de indivíduos de cada sexo em cada grupo etário e nível de escolaridade, no final de cada quinquénio de projeção. Para tal, recorreremos a uma matriz de Leslie em blocos, designada por matriz \mathbf{B} , que inclui as proporções de sobrevivência, taxas de fertilidade e probabilidades de transição, e a um vetor correspondente ao saldo migratório. Ficamos com a seguinte expressão, para cada um dos sexos:

$$\ell_{t+5} = \mathbf{B}\ell_t + \mathbf{SM}_t$$

Tomando como exemplo a população do sexo feminino, representada através do expoente f , temos:

$$\begin{bmatrix} \ell_{1,1,t+5}^f \\ \ell_{1,2,t+5}^f \\ \ell_{1,3,t+5}^f \\ \ell_{1,4,t+5}^f \\ \ell_{2,1,t+5}^f \\ \ell_{2,2,t+5}^f \\ \vdots \\ \vdots \\ \vdots \\ \ell_{m,4,t+5}^f \end{bmatrix} = \begin{bmatrix} r \cdot b_{11} \cdot p_{01} & r \cdot b_{12} \cdot p_{02} & r \cdot b_{13} \cdot p_{03} & r \cdot b_{14} \cdot p_{04} & r \cdot b_{21} \cdot p_{01} & r \cdot b_{22} \cdot p_{02} & r \cdot b_{23} \cdot p_{03} & r \cdot b_{24} \cdot p_{04} & \dots & r \cdot b_{m1} \cdot p_{01} & r \cdot b_{m2} \cdot p_{02} & r \cdot b_{m3} \cdot p_{03} & r \cdot b_{m4} \cdot p_{04} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ p_{111} & p_{121} & p_{131} & p_{141} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ p_{112} & p_{122} & p_{132} & p_{142} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ p_{113} & p_{123} & p_{133} & p_{143} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ p_{114} & p_{124} & p_{134} & p_{144} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{211} & p_{221} & p_{231} & p_{241} & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{212} & p_{222} & p_{232} & p_{242} & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{213} & p_{223} & p_{233} & p_{243} & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{214} & p_{224} & p_{234} & p_{244} & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \ell_{1,1,t}^f \\ \ell_{1,2,t}^f \\ \ell_{1,3,t}^f \\ \ell_{1,4,t}^f \\ \ell_{2,1,t}^f \\ \ell_{2,2,t}^f \\ \vdots \\ \vdots \\ \vdots \\ \ell_{m,4,t}^f \end{bmatrix} + \begin{bmatrix} SM_{1,1,t}^f \\ SM_{1,2,t}^f \\ SM_{1,3,t}^f \\ SM_{1,4,t}^f \\ SM_{2,1,t}^f \\ SM_{2,2,t}^f \\ \vdots \\ \vdots \\ \vdots \\ SM_{m,4,t}^f \end{bmatrix}$$

onde ℓ_{ijt}^f representa o número de mulheres no grupo etário i e estado j em t ; b_{ij} representa a taxa de fecundidade específica para as mulheres do grupo etário i e estado j (notemos que somente a primeira linha da tabela inclui a contribuição das mulheres para a formação do primeiro grupo etário, dado que todos os nascimentos irão pertencer ao primeiro estado, Sem escolaridade/2.º Ciclo EB); r indica probabilidade dos nascimentos ocorridos no intervalo serem do sexo feminino; p_{ijk} refere-se à probabilidade dos indivíduos que, em t , estão no grupo etário i e estado j sobreviverem, encontrando-se no estado k em $t+5$; e SM_{ijt}^f indica o saldo migratório entre os indivíduos do sexo feminino, grupo etário i e estado j , durante o período $[t, t+5[$. A matriz de projeção \mathbf{B} pode ser ainda simplificada, com base nos pressupostos assumidos anteriormente.

Ao considerarmos que as mulheres em idade reprodutiva têm idades compreendidas entre os 15 e os 49 anos e, como tal, as taxas de fecundidade específicas para os grupos etários fora deste intervalo são nulas, os únicos valores não nulos na primeira linha da matriz \mathbf{B} serão os relativos aos grupos etários $i = 4, 5, \dots, 10$. Para além disso, sabemos que as transições para um nível de escolaridade inferior são impossíveis, seja qual for o grupo etário considerado, e, em consequência, o valor de p_{ijk} será nulo sempre que $j > k$. Também as transições dos estados Sem escolaridade/2.ºCiclo EB e Ensino Básico para o Ensino Superior são consideradas impossíveis no intervalo de tempo em que trabalhamos (5 anos). Focando-nos depois em cada grupo etário específico, verificamos que, consoante os eventos que puderem ocorrer nessa idade, tal como assumimos anteriormente, existirão outras probabilidades de transição cujo valor será necessariamente nulo: por exemplo, admitimos que os indivíduos com idades entre os 20 e os 24 anos em t e que terão, portanto, 25 a 29 anos em $t+5$, podem transitar do Ensino Secundário para o Ensino Superior nesse intervalo de 5 anos, mas já não farão nenhuma das transições anteriores. Nesse sentido, quem tem, em t , o Ensino Básico no grupo etário 20-24 anos, permanecerá nesse nível de escolaridade no final do quinquénio seguinte. Assim, as probabilidades p_{ijk} transformam-se em probabilidades de sobrevivência simples, nos casos em que já não existem transições possíveis entre j e k : passam a ser designadas por p_{ij} e representam a probabilidade das pessoas no grupo etário i e estado j sobreviverem os 5 anos, passando para o grupo etário $i+1$, mas mantendo-se no estado j (estas probabilidades serão estimadas com base nas proporções de sobrevivência apresentadas na secção 5.2.3.1).

As probabilidades dos indivíduos sobreviverem e permanecerem num dado estado, nos grupos etários em que ocorrem transições, são dadas por p_{ijj} , e são obtidas como $p_{ijj} = p_{ij} - \sum_{k \neq j} p_{ijk}$. Representando, para os elementos de um dado grupo etário i no nível de escolaridade j , em t , a probabilidade de morte no intervalo $[t, t + 5[$ por $1 - p_{ij}$, a probabilidade de sobrevivência e transição para qualquer outro nível de escolaridade por $\sum_{k \neq j} p_{ijk}$ e a probabilidade de sobrevivência e manutenção no mesmo nível por p_{ijj} , teremos que $(1 - p_{ij}) + p_{ijj} + \sum_{k \neq j} p_{ijk} = 1 \Leftrightarrow p_{ijj} = p_{ij} - \sum_{k \neq j} p_{ijk}$.

A projeção da população do sexo masculino (distinguida pelo expoente m) efetua-se de forma semelhante, mas, neste caso, a matriz \mathbf{B} não inclui taxas de fecundidade e o total de nascimentos em $[t, t + 5[$ é dado por:

$$\ell_{1,1,t+5}^m = (1 - r) \sum_{i=4}^{10} \sum_{j=1}^4 b_{ij} \ell_{ijt}^f$$

A projeção da população para o instante correspondente a 1 de janeiro de 2006 tem por base o vetor ℓ_{2001} correspondente à população de partida (ver secção 5.1) e as estimativas das probabilidades de sobrevivência²⁵ e de transição, das taxas de fecundidade e dos valores de saldo migratório obtidas com base nos passos apresentados anteriormente (ver apêndice C). A população projetada para o final de cada quinquénio é utilizada como população inicial no quinquénio seguinte e os parâmetros relativos aos intervalos posteriores a 2006 são obtidos de acordo com a metodologia apresentada na secção seguinte.

5.3. Tratamento da incerteza: construção de cenários de projeção

5.3.1. Mortalidade, fecundidade e migrações. Analogamente à metodologia seguida por Lutz et al. (2005) e KC et al. (2010), define-se somente um cenário relativo à evolução das variáveis demográficas fecundidade, mortalidade e migrações.

A obtenção de valores relativos às taxas de fecundidade específicas por idade e nível de escolaridade para os quinquénios seguintes a 2006 baseia-se nos resultados da mais recente revisão das estimativas e projeções demográficas oficiais da Divisão de População das Nações Unidas (ONU, 2011b). Estas projeções, referentes ao período 2010-2100, incluem um conjunto de variantes relativas a cada uma das componentes, traduzidas posteriormente em oito cenários. Para projetar a fecundidade é utilizado um modelo estocástico, baseado em tendências empíricas estimadas para o período de 1950 a 2010, que decompõe a evolução do Índice Sintético de Fecundidade em três fases: (i) elevada fecundidade pré-transição, (ii) transição de fecundidade, e (iii) baixa fecundidade pós-transição. A diminuição do ISF correspondente à fase de transição (fase II), em que Portugal se encontrava no momento de elaboração

²⁵Para projetar o número de indivíduos no último grupo etário (75 ou mais anos) assume-se que tanto os indivíduos com 70-74 anos como aqueles que já têm 75 ou mais anos no instante t , são sujeitos às proporções de sobrevivência calculadas para o penúltimo grupo etário.

das projeções²⁶, é modelada como a soma de duas funções logísticas que dependem do valor atual de ISF e de um termo aleatório, através de um modelo bayesiano hierárquico. Este modelo é utilizado para projetar os valores futuros de ISF com base tanto nos valores observados no passado em cada país, como no padrão de todos os países, e é estimado através de um algoritmo de Monte Carlo via Cadeias de Markov (MCMC). A fase de baixa fertilidade pós-transição é modelada com recurso a um modelo autoregressivo, segundo o qual as projeções de longo prazo para o ISF convergem para o nível mínimo de substituição de gerações (2,1 filhos por mulher em idade reprodutiva) e oscilam em torno deste valor (Alkema et al., 2011). O pressuposto de que exista, no longo prazo, uma recuperação dos níveis de fecundidade nos países que completam a transição, para níveis em torno do nível mínimo de substituição, é fundamentado na experiência de muitos países da Europa e Leste da Ásia com níveis de fecundidade abaixo do nível de substituição (ONU, 2011a). A variante média para a fecundidade, tida como aquela que representa o cenário futuro mais provável, é dada pela mediana de 100.000 trajetórias para o ISF projetadas para cada país. A obtenção de taxas de fecundidade específicas por idade e nível de escolaridade para os quinquênios de projeção posteriores a 2006 baseia-se nas taxas de fertilidade específicas por idade referentes a esta variante (tabela A.0.12, apêndice A), assumindo-se que se mantém, até ao final do período de projeção, os diferenciais de fecundidade por nível de escolaridade verificados, em cada grupo etário, no quinquénio 2001-2006 (obtidos com base nas taxas de fecundidade específicas calculadas na secção 5.2.3.2). Os valores publicados pelas Nações Unidas referentes ao quinquénio 2005-2010 servem de base à estimação das taxas relativas ao período 2006-2011, os que correspondem ao período 2010-2015 são usados para estimar as taxas referentes a 2011-2016, e assim sucessivamente. As taxas de fecundidade resultantes são apresentadas na tabela 5.3.2. De acordo com estes valores estimados, os níveis de fecundidade das mulheres nos grupos etários mais jovens (15-19 e 20-24 anos) diminuirão significativamente ao longo do período de projeção. O mesmo acontecerá em relação às mulheres mais velhas (45-49 anos), que apresentam, desde o início do período, os mais baixos níveis de fecundidade. Pelo contrário, as mulheres entre os 25 e os 39 anos apresentarão taxas de fecundidade cada vez mais elevadas, à medida que avançamos até 2031. Ocorrerá um adiamento da fecundidade, com maior evidência entre as mulheres que não completaram o Ensino Básico: neste grupo, os maiores níveis de fecundidade ocorrerão entre os 20 e os 24 anos até 2016-2021 e, daí em diante, entre os 25 e os 29 anos. Nos restantes níveis de escolaridade, as idades em que se verificam maiores níveis de fecundidade manter-se-ão ao longo do tempo (entre os 25 e os 29 anos para as mulheres com Ensino Básico e entre os 30 e os 34 anos para as que completaram pelo menos o Ensino Secundário). Verifica-se, portanto, que as mulheres com níveis de escolaridade mais elevados continuam a ser as que têm os seus filhos num momento mais tardio da sua vida; ainda assim, considerando todas as mulheres em idade reprodutiva, são as que completaram algum nível de Ensino Superior que apresentam, ao longo de todo o período, os níveis de fecundidade mais

²⁶De acordo com Alkema et al. (2011), considera-se que o início da fase III ocorre quando se observam dois aumentos sucessivos do ISF, mantendo-se este indicador abaixo dos 2 filhos por mulher. De acordo com esta definição, 20 países transitaram para a fase III em 2005-2010: Bélgica, Bulgária, Ilhas do Canal, República Checa, Dinamarca, Estónia, Finlândia, França, Alemanha, Itália, Letónia, Luxemburgo, Países Baixos, Noruega, Federação Russa, Singapura, Espanha, Suécia, Reino Unido e Estados Unidos da América.

elevados. De acordo com os Índices Sintéticos de Fecundidade para cada grupo de escolaridade, em cada quinquénio (tabela 5.3.1), verificamos ainda que são as mulheres com menor grau de instrução que terão, em média, menos filhos.

TABELA 5.3.1. *Índice Sintético de Fecundidade, por nível de escolaridade, 2001-2031*

Fonte: Cálculos do autor

Quinquénio	SE/2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior
2001-2006	1,56	1,46	1,53	1,64
2006-2011	1,45	1,40	1,49	1,63
2011-2016	1,34	1,35	1,46	1,63
2016-2021	1,27	1,34	1,47	1,67
2021-2026	1,27	1,40	1,56	1,80
2026-2031	1,28	1,48	1,67	1,97

TABELA 5.3.2. Taxa de Fecundidade (‰) por grupo etário e nível de escolaridade da mãe, 2001-2031

Fonte: Cálculos do autor

Quinquênio	Grupo etário	SE/2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior
2001-2006	15-19	38,189	11,863	16,540	N.A.
	20-24	109,792	60,631	28,381	18,420
	25-29	82,450	98,455	93,466	77,363
	30-34	52,666	78,897	102,634	145,784
	35-39	22,966	35,429	53,247	70,514
	40-44	5,080	6,908	10,772	14,448
	45-49	0,325	0,495	0,728	0,872
2006-2011	15-19	32,770	10,179	14,193	N.A.
	20-24	99,178	54,770	25,637	16,639
	25-29	74,741	89,250	84,727	70,129
	30-34	52,937	79,304	103,163	146,536
	35-39	25,095	38,713	58,184	77,051
	40-44	5,432	7,387	11,518	15,449
	45-49	0,284	0,431	0,635	0,761
2011-2016	15-19	24,450	7,595	10,590	0,000
	20-24	85,585	47,263	22,124	14,359
	25-29	73,703	88,010	83,550	69,155
	30-34	53,691	80,433	104,632	148,623
	35-39	25,564	39,437	59,271	78,491
	40-44	5,118	6,960	10,853	14,556
	45-49	0,216	0,329	0,485	0,581
2016-2021	15-19	17,225	5,351	7,460	N.A.
	20-24	75,043	41,441	19,398	12,590
	25-29	74,779	89,295	84,770	70,165
	30-34	55,959	83,831	109,052	154,901
	35-39	26,730	41,236	61,975	82,071
	40-44	4,939	6,717	10,474	14,048
	45-49	0,172	0,261	0,384	0,460
2021-2026	15-19	10,604	3,294	4,593	N.A.
	20-24	68,181	37,652	17,625	11,439
	25-29	79,892	95,401	90,567	74,963
	30-34	61,262	91,775	119,386	169,579
	35-39	29,353	45,283	68,058	90,127
	40-44	5,039	6,853	10,686	14,332
	45-49	0,119	0,182	0,267	0,320
2026-2031	15-19	3,476	1,080	1,506	N.A.
	20-24	61,167	33,779	15,811	10,262
	25-29	86,272	103,020	97,799	80,949
	30-34	67,657	101,356	131,850	187,283
	35-39	32,516	50,161	75,390	99,836
	40-44	5,182	7,047	10,989	14,738
	45-49	0,067	0,102	0,150	0,180

N.A.: Não aplicável

No que diz respeito à mortalidade, os valores de esperança de vida à nascença e taxa de mortalidade infantil específicos para cada sexo, relativos também à variante média das projeções das Nações Unidas (tabelas A.0.13 e A.0.14, apêndice A), foram utilizados para estimar tabelas de mortalidade para cada quinquénio de projeção. Este procedimento baseia-se na metodologia apresentada por Clark e Sharrow (2011), que utilizaram técnicas de análise classificatória para identificar cinco padrões etários de mortalidade em 844 tabelas de mortalidade da *Human Mortality Database* (California e Demographic Research, 2009) e, utilizando cada padrão como base para uma "família" de um sistema tradicional de tabelas-tipo de mortalidade²⁷, criaram um modelo que distingue níveis de mortalidade em cada família. Em consequência, obtiveram um sistema de tabelas-tipo com dois parâmetros, organizado em famílias, com vários níveis de mortalidade dentro de cada família, para os países e períodos incluídos na *Human Mortality Database*²⁸. O programa *LifeTables* do *software* R permite, através do recurso a análise discriminante, identificar a família mais apropriada, com base em uma ou mais medidas de mortalidade adulta ou infantil fornecidas pelo utilizador. Para além disso, o utilizador pode escolher um nível de esperança de vida dentro da família identificada e, assim, obter uma tabela-tipo de mortalidade baseada na família e no nível selecionado (Sharrow e Sevcikova, 2011). Obtidas as tabelas de mortalidade específicas para o sexo, a partir dos valores de esperança de vida à nascença e taxa de mortalidade infantil correspondentes à variante média das Nações Unidas, recorreu-se à metodologia aplicada para o primeiro quinquénio de projeção (secção 5.2.3.1), para estimar as proporções de sobrevivência específicas por sexo, grupo etário e nível de escolaridade, para os quinquénios seguintes. A média dos diferenciais das taxas de mortalidade infantil observadas para o período 2005-2010, por nível de escolaridade da mãe, foi utilizada para introduzir os diferenciais de mortalidade por grau de instrução nos restantes grupos etários, em todos os quinquénios subsequentes a 2006. As proporções de sobrevivência obtidas são apresentadas na tabela 5.3.3. Coincidindo com o comportamento observado nos últimos anos, as proporções estimadas apontam para a existência de um diferencial no que respeita à mortalidade entre os sexos: as mulheres tendem a apresentar proporções de sobrevivência mais elevadas em relação aos homens, para o mesmo grupo etário e nível de escolaridade, importando destacar que esta diferença tende a diminuir à medida que avançamos no tempo de calendário. Como é de esperar, as proporções de sobrevivência diminuem sempre com a idade, seja qual for o sexo, nível de escolaridade e quinquénio considerados. Por outro lado, os níveis de sobrevivência deverão continuar a ser superiores para indivíduos com níveis de escolaridade mais elevados, para o mesmo sexo, grupo etário e quinquénio.

²⁷A construção de sistemas de tabelas-tipo de mortalidade baseia-se nas semelhanças observadas entre padrões etários de mortalidade de diferentes populações, existindo atualmente um conjunto de modelos com diferentes graus de complexidade e sofisticação. Os sistemas mais conhecidos são as tabelas-tipo de Coale-Demeny (Coale, 1966), as tabelas-tipo das Nações Unidas para os países em desenvolvimento (UN, 1982), as tabelas-tipo da Organização Mundial de Saúde (Murray et al., 2000), o sistema de Ledermann (Ledermann, 1969) e o sistema de Brass Logit (Brass, 1971). Dado que estes sistemas são gerados com base num conjunto de tabelas de sobrevivência empíricas, apenas podem refletir a variação nos padrões de mortalidade contida nos dados utilizados.

²⁸Tendo em conta que os sistemas mais utilizados foram construídos há várias décadas - por exemplo, o sistema de Coale-Demeny (Coale, 1966) não permite considerar níveis de mortalidade infantil extremamente baixos, pois este tipo de perfil de mortalidade não existia na altura em que foi produzido - optamos por recorrer antes a um sistema baseado em dados contemporâneos.

TABELA 5.3.3. *Proporções de sobrevivência por sexo, grupo etário e nível de escolaridade, 2001-2031*

Fonte: Cálculos do autor

Quinquênio	Grupo etário	H				M			
		SE/2.* Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	SE/2.* Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior
2001-2006	0	0,994	N.A.	N.A.	N.A.	0,995	N.A.	N.A.	N.A.
	1-4	0,998	N.A.	N.A.	N.A.	0,998	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	10-14	0,997	0,997	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,995	0,995	0,997	N.A.	0,998	0,998	0,999	N.A.
	20-24	0,993	0,993	0,995	0,997	0,998	0,998	0,999	0,999
	25-29	0,991	0,991	0,994	0,997	0,997	0,997	0,998	0,999
	30-34	0,987	0,988	0,992	0,995	0,996	0,996	0,997	0,998
	35-39	0,982	0,983	0,989	0,994	0,993	0,994	0,996	0,998
	40-44	0,975	0,977	0,984	0,991	0,990	0,991	0,994	0,996
	45-49	0,966	0,969	0,978	0,988	0,986	0,987	0,991	0,995
	50-54	0,954	0,957	0,970	0,983	0,980	0,981	0,987	0,993
	55-59	0,933	0,937	0,957	0,976	0,970	0,972	0,981	0,989
	60-64	0,898	0,905	0,934	0,963	0,952	0,955	0,969	0,983
	65-69	0,839	0,849	0,896	0,941	0,916	0,922	0,946	0,970
	70-74	0,578	0,593	0,666	0,747	0,646	0,657	0,710	0,770
2006-2011	0	0,992	N.A.	N.A.	N.A.	0,995	N.A.	N.A.	N.A.
	1-4	0,998	N.A.	N.A.	N.A.	0,998	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	10-14	0,997	0,998	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,994	0,995	0,996	N.A.	0,998	0,999	0,999	N.A.
	20-24	0,993	0,994	0,996	0,997	0,998	0,998	0,999	0,999
	25-29	0,993	0,994	0,995	0,997	0,997	0,998	0,998	0,999
	30-34	0,991	0,993	0,994	0,997	0,996	0,997	0,998	0,999
	35-39	0,988	0,990	0,992	0,995	0,995	0,995	0,997	0,998
	40-44	0,981	0,985	0,988	0,993	0,991	0,993	0,994	0,997
	45-49	0,971	0,976	0,981	0,989	0,986	0,988	0,991	0,995
	50-54	0,954	0,962	0,971	0,982	0,978	0,982	0,986	0,992
	55-59	0,928	0,940	0,954	0,972	0,967	0,972	0,979	0,987
	60-64	0,887	0,906	0,928	0,955	0,946	0,955	0,966	0,979
	65-69	0,824	0,853	0,887	0,930	0,910	0,925	0,942	0,964
	70-74	0,576	0,618	0,670	0,738	0,656	0,685	0,721	0,768
2011-2016	0	0,994	N.A.	N.A.	N.A.	0,996	N.A.	N.A.	N.A.
	1-4	0,998	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	10-14	0,998	0,998	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,995	0,996	0,997	N.A.	0,999	0,999	0,999	N.A.
	20-24	0,994	0,995	0,996	0,998	0,998	0,999	0,999	0,999
	25-29	0,994	0,995	0,996	0,998	0,998	0,998	0,999	0,999
	30-34	0,993	0,994	0,995	0,997	0,997	0,997	0,998	0,999
	35-39	0,990	0,991	0,993	0,996	0,995	0,996	0,997	0,998
	40-44	0,984	0,987	0,990	0,994	0,992	0,994	0,995	0,997
	45-49	0,975	0,979	0,984	0,990	0,988	0,990	0,992	0,995
	50-54	0,960	0,967	0,974	0,984	0,981	0,984	0,988	0,992
	55-59	0,936	0,947	0,959	0,975	0,970	0,975	0,981	0,988
	60-64	0,899	0,917	0,936	0,960	0,952	0,960	0,969	0,981
	65-69	0,842	0,869	0,900	0,938	0,918	0,932	0,948	0,968
	70-74	0,597	0,636	0,684	0,746	0,668	0,695	0,729	0,772
2016-2021	0	0,995	N.A.	N.A.	N.A.	0,996	N.A.	N.A.	N.A.
	1-4	0,998	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	10-14	0,998	0,998	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,996	0,996	0,997	N.A.	0,999	0,999	0,999	N.A.
	20-24	0,995	0,996	0,997	0,998	0,998	0,999	0,999	0,999
	25-29	0,994	0,995	0,996	0,998	0,998	0,998	0,999	0,999
	30-34	0,993	0,994	0,996	0,997	0,997	0,998	0,998	0,999
	35-39	0,990	0,992	0,994	0,996	0,996	0,996	0,997	0,998
	40-44	0,985	0,988	0,991	0,994	0,993	0,994	0,995	0,997
	45-49	0,976	0,980	0,985	0,991	0,988	0,990	0,993	0,995
	50-54	0,962	0,968	0,976	0,985	0,982	0,985	0,988	0,993
	55-59	0,939	0,950	0,962	0,976	0,972	0,977	0,982	0,989
	60-64	0,904	0,921	0,939	0,962	0,955	0,962	0,971	0,982
	65-69	0,849	0,875	0,904	0,941	0,923	0,936	0,951	0,970
	70-74	0,605	0,643	0,690	0,749	0,674	0,700	0,733	0,775
2021-2026	0	0,999	N.A.	N.A.	N.A.	0,997	N.A.	N.A.	N.A.
	1-4	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	1,000	N.A.	N.A.	N.A.
	10-14	0,998	0,998	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,996	0,997	0,997	N.A.	0,999	0,999	0,999	N.A.
	20-24	0,995	0,996	0,997	0,998	0,999	0,999	0,999	0,999
	25-29	0,995	0,996	0,997	0,998	0,998	0,998	0,999	0,999
	30-34	0,994	0,995	0,996	0,998	0,997	0,998	0,998	0,999
	35-39	0,991	0,993	0,994	0,997	0,996	0,997	0,997	0,998
	40-44	0,986	0,988	0,991	0,995	0,993	0,994	0,996	0,997
	45-49	0,978	0,981	0,986	0,991	0,989	0,991	0,993	0,996
	50-54	0,964	0,970	0,977	0,986	0,983	0,986	0,989	0,993
	55-59	0,943	0,952	0,964	0,977	0,974	0,978	0,983	0,990
	60-64	0,909	0,924	0,942	0,964	0,957	0,965	0,973	0,983
	65-69	0,856	0,881	0,909	0,943	0,927	0,939	0,954	0,971
	70-74	0,613	0,650	0,695	0,753	0,680	0,706	0,737	0,777
2026-2031	0	0,996	N.A.	N.A.	N.A.	0,997	N.A.	N.A.	N.A.
	1-4	0,999	N.A.	N.A.	N.A.	0,999	N.A.	N.A.	N.A.
	5-9	0,999	N.A.	N.A.	N.A.	1,000	N.A.	N.A.	N.A.
	10-14	0,998	0,998	N.A.	N.A.	0,999	0,999	N.A.	N.A.
	15-19	0,996	0,997	0,998	N.A.	0,999	0,999	0,999	N.A.
	20-24	0,995	0,996	0,997	0,998	0,999	0,999	0,999	0,999
	25-29	0,995	0,996	0,997	0,998	0,998	0,999	0,999	0,999
	30-34	0,994	0,995	0,996	0,998	0,998	0,998	0,998	0,999
	35-39	0,992	0,993	0,995	0,997	0,996	0,997	0,998	0,999
	40-44	0,987	0,989	0,992	0,995	0,994	0,995	0,996	0,998
	45-49	0,979	0,982	0,987	0,992	0,990	0,992	0,994	0,996
	50-54	0,966	0,972	0,978	0,987	0,984	0,987	0,990	0,994
	55-59	0,945	0,955	0,965	0,979	0,975	0,979	0,984	0,990
	60-64	0,913	0,928	0,945	0,966	0,960	0,967	0,974	0,984
	65-69	0,862	0,886	0,912	0,946	0,931	0,943	0,956	0,973
	70-74	0,620	0,656	0,700	0,755	0,685	0,710	0,741	0,779

N.A.: Não aplicável

Os valores de saldo migratório quinquenal por sexo, grupo etário e nível de escolaridade até ao final do período de projeção foram calculados com recurso à metodologia aplicada para o primeiro quinquénio (secção 5.2.3.3). Para tal, foram utilizadas as estimativas anuais de saldo migratório (Estatísticas Demográficas do INE, I.P.) até ao último período disponível (2010) e, daí em diante e até 2031, os saldos migratórios anuais relativos ao cenário central das projeções demográficas elaboradas pelo INE, I.P., para o período 2008-2060^{29,30} (o conjunto de dados utilizados encontra-se na tabela A.0.15, apêndice A). A distribuição do saldo migratório total por sexo e grupo etário teve por base a estrutura de movimentos migratórios internacionais para Portugal (UN, 2011) relativa ao ano 2010 e, para a posterior distribuição dos indivíduos de cada coorte pelos vários níveis de escolaridade, foi considerada a média das proporções de indivíduos em cada nível entre os anos 2006 e 2010. Os valores obtidos são apresentados na tabela 5.3.4.

²⁹Os resultados das projeções de população residente em Portugal elaboradas pelo INE, I.P., incluem quatro cenários, que derivam de diferentes combinações das variantes consideradas para a evolução de cada uma das componentes. Perante a evolução demográfica recente, é o cenário central que combina o conjunto de hipóteses considerado como sendo o mais provável (INE, 2009). No que respeita às migrações, com base na análise dos valores da média dos saldos migratórios estimados para os 15 anos anteriores ao início do período de projeção, pressupõe-se, no cenário central, um aumento ligeiro do saldo migratório anual até atingir os 36,6 mil indivíduos no ano de 2018, momento a partir do qual o mesmo se mantém constante até ao final do período de projeção.

³⁰Para a obtenção de valores futuros relativos às migrações, optou-se por utilizar dados resultantes das projeções do INE, I.P., que, por serem anuais, permitem aplicar a mesma metodologia que foi usada para o primeiro quinquénio, ao contrário dos dados das projeções das Nações Unidas, que apenas são disponibilizados em relação a períodos quinquenais (que não coincidem exatamente com os quinquénios considerados neste trabalho); por outro lado, não existem resultados das projeções do INE, I.P., relativos às taxas de fecundidade específicas por idade nem às taxas de mortalidade infantil, o que conduziu à necessidade de recorrer às projeções das Nações Unidas.

TABELA 5.3.4. Saldo migratório por sexo, grupo etário e nível de escolaridade, 2001-2031

Fonte: Cálculos do autor

Quinquênio	Grupo etário	H				M			
		SE/2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	SE/2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior
2001-2006	0-4	2186	N.A.	N.A.	N.A.	2121	N.A.	N.A.	N.A.
	5-9	4656	N.A.	N.A.	N.A.	4527	N.A.	N.A.	N.A.
	10-14	6389	101	N.A.	N.A.	6208	102	N.A.	N.A.
	15-19	3686	4963	748	N.A.	2534	5460	1312	N.A.
	20-24	4028	4728	5016	429	2643	3607	6454	1282
	25-29	9183	5217	5654	3052	7763	3632	5755	5499
	30-34	10193	3313	3014	2006	8683	2566	3215	3304
	35-39	8525	2273	1768	1274	7681	1865	1769	1994
	40-44	7036	1604	1215	875	6696	1540	1200	1205
	45-49	5258	944	727	619	5726	922	664	835
	50-54	3502	523	357	330	4184	478	291	440
	55-59	2517	328	214	221	3247	283	156	251
	60-64	2155	200	112	134	3181	163	98	167
	65-69	1794	131	73	77	2556	96	59	88
	70-74	1477	88	52	58	2281	71	46	56
	75 +	2052	98	60	76	3925	100	61	67
2006-2011	0-4	788	N.A.	N.A.	N.A.	754	N.A.	N.A.	N.A.
	5-9	1324	N.A.	N.A.	N.A.	1214	N.A.	N.A.	N.A.
	10-14	1629	26	N.A.	N.A.	1566	26	N.A.	N.A.
	15-19	843	1399	282	N.A.	549	1507	414	N.A.
	20-24	725	1347	1678	222	470	1030	1860	451
	25-29	1435	1697	1767	1058	1096	1190	1646	1721
	30-34	2506	1499	1273	968	1787	1101	1421	1553
	35-39	2569	1034	961	689	2082	837	960	1021
	40-44	2363	762	511	419	2054	681	551	608
	45-49	1885	526	345	265	1780	562	376	380
	50-54	1287	307	203	189	1467	297	203	253
	55-59	1006	164	110	127	1292	173	98	155
	60-64	788	107	62	78	1120	110	61	92
	65-69	616	57	33	37	872	46	32	48
	70-74	507	39	24	28	778	34	25	31
	75 +	705	43	27	36	1339	48	34	36
2011-2016	0-4	1397	N.A.	N.A.	N.A.	1340	N.A.	N.A.	N.A.
	5-9	2353	N.A.	N.A.	N.A.	2158	N.A.	N.A.	N.A.
	10-14	2896	46	N.A.	N.A.	2783	46	N.A.	N.A.
	15-19	1499	2486	501	N.A.	976	2679	737	N.A.
	20-24	1288	2395	2982	394	836	1831	3305	801
	25-29	2551	3017	3142	1880	1948	2115	2926	3059
	30-34	4455	2665	2263	1721	3176	1957	2526	2761
	35-39	4567	1839	1708	1225	3700	1488	1706	1815
	40-44	4201	1354	908	745	3652	1211	979	1080
	45-49	3350	934	614	472	3164	999	669	676
	50-54	2289	545	361	336	2607	527	360	450
	55-59	1788	291	195	226	2296	307	173	275
	60-64	1400	190	109	138	1990	195	108	163
	65-69	1095	102	59	65	1549	81	49	85
	70-74	902	69	42	49	1383	60	38	54
	75 +	1253	77	49	64	2379	85	52	65
2016-2021	0-4	1691	N.A.	N.A.	N.A.	1621	N.A.	N.A.	N.A.
	5-9	2947	55	N.A.	N.A.	2612	N.A.	N.A.	N.A.
	10-14	3504	55	N.A.	N.A.	3367	55	N.A.	N.A.
	15-19	1814	3008	606	N.A.	1181	3241	891	N.A.
	20-24	1559	2898	3606	477	1011	2216	3999	969
	25-29	3087	3651	3801	2275	2357	2559	3540	3701
	30-34	5390	3224	2738	2082	3642	2368	3057	3340
	35-39	5526	2225	2067	1482	4477	1801	2064	2196
	40-44	5083	1638	1098	901	4418	1465	1184	1307
	45-49	4054	1130	743	579	3628	1208	809	808
	50-54	2769	660	436	403	3154	638	436	525
	55-59	2163	353	236	265	2779	371	210	318
	60-64	1694	230	132	156	2408	236	131	181
	65-69	1325	123	71	124	1875	98	60	207
	70-74	1091	83	51	94	1673	73	46	133
	75 +	1516	93	59	122	2879	102	62	158
2021-2026	0-4	1727	N.A.	N.A.	N.A.	1656	N.A.	N.A.	N.A.
	5-9	2909	N.A.	N.A.	N.A.	2668	N.A.	N.A.	N.A.
	10-14	3580	56	N.A.	N.A.	3440	56	N.A.	N.A.
	15-19	1853	3073	619	N.A.	1206	3312	911	N.A.
	20-24	1593	2961	3686	488	1033	2264	4086	990
	25-29	3154	3730	3884	2324	2408	2615	3617	3782
	30-34	5507	3294	2798	2128	3926	2419	3123	3413
	35-39	5645	2273	2112	1514	4574	1840	2109	2243
	40-44	5193	1673	1122	921	4514	1497	1210	1335
	45-49	4141	1155	759	583	3911	1234	827	835
	50-54	2829	674	446	415	3223	652	445	556
	55-59	2210	360	241	279	2839	379	214	340
	60-64	1731	235	135	171	2460	241	134	202
	65-69	1353	126	73	80	1915	100	61	105
	70-74	1115	85	52	61	1709	74	47	67
	75 +	1549	95	60	80	2941	105	64	80
2026-2031	0-4	1727	N.A.	N.A.	N.A.	1656	N.A.	N.A.	N.A.
	5-9	2909	N.A.	N.A.	N.A.	2668	N.A.	N.A.	N.A.
	10-14	3580	56	N.A.	N.A.	3440	56	N.A.	N.A.
	15-19	1853	3073	619	N.A.	1206	3312	911	N.A.
	20-24	1593	2961	3686	488	1033	2264	4086	990
	25-29	3154	3730	3884	2324	2408	2615	3617	3782
	30-34	5507	3294	2798	2128	3926	2419	3123	3413
	35-39	5645	2273	2112	1514	4574	1840	2109	2243
	40-44	5193	1673	1122	921	4514	1497	1210	1335
	45-49	4141	1155	759	583	3911	1234	827	835
	50-54	2829	674	446	415	3223	652	445	556
	55-59	2210	360	241	279	2839	379	214	340
	60-64	1731	235	135	171	2460	241	134	202
	65-69	1353	126	73	80	1915	100	61	105
	70-74	1115	85	52	61	1709	74	47	67
	75 +	1549	95	60	80	2941	105	64	80

N.A.: Não aplicável

5.3.2. Escolarização. A elaboração de hipóteses acerca do desenvolvimento futuro da escolaridade da população deve ser visto, tal como acontece em relação à fecundidade, mortalidade ou migrações, e porque nos focamos na elaboração de projeções determinísticas, como um exercício com base num raciocínio do tipo “e se...”. Dado que estas componentes da dinâmica demográfica são influenciadas pelos níveis de escolaridade, KC et al. (2010) defendem que as projeções populacionais assumem sempre, implicitamente, hipóteses sobre o perfil educacional futuro da população, mesmo que estas hipóteses não sejam declaradas com clareza e que, portanto, é preferível assumi-las explicitamente.

5.3.2.1. *Cenário constante.* Partindo das probabilidades de transição calculadas anteriormente, relativas ao quinquénio 2001-2006, pode ser admitido um cenário de escolarização constante, em que as mesmas se mantêm ao longo do período de projeção. Este cenário resulta em proporções de indivíduos em cada um dos níveis de escolaridade constantes para as várias coortes e, por isso, serve apenas o propósito de demonstrar o efeito de estender para o futuro o contexto inicial, sem que seja levada em conta a tendência recente nem o efeito de eventuais alterações das circunstâncias de partida.

5.3.2.2. *Cenário de tendência.* Alternativamente, podemos considerar um cenário de tendência, de acordo com o qual se assume que os níveis de escolarização continuarão, no futuro, a seguir a tendência verificada nos últimos anos. Neste sentido, é habitualmente aplicado um de dois métodos alternativos para estimar as proporções futuras de indivíduos em cada nível de escolaridade: (1) modelar a tendência das proporções de transição, geralmente com recurso às taxas de escolarização de coortes sequenciais, e utilizar as proporções de transição projetadas para estimar as proporções de estado no futuro; ou (2) modelar a tendência das proporções de estado, de forma a obter estimativas das proporções de indivíduos em cada nível de escolaridade nos instantes futuros e, com base nestas, calcular as proporções de transição correspondentes. Na primeira abordagem, seguida por Lutz et al. (2005), Wils et al. (2005) e Wils (2007), recorre-se à proporção de indivíduos por nível de escolaridade em cada grupo etário, no ano de partida, para obter uma estimativa do padrão de escolarização no momento em que esse grupo de indivíduos tinha a idade prevista para conclusão de cada nível (por exemplo, a proporção de indivíduos com o Ensino Básico em cada grupo etário, no ano de 2001, pode ser utilizada como uma estimativa da percentagem de indivíduos que atingiram este nível de escolaridade quando cada coorte tinha 15 anos³¹, sendo o mesmo raciocínio aplicado aos restantes níveis de ensino). Como resultado, a tendência pode ser aplicada ao grupo etário em que dada transição ocorre com maior frequência, mas não às restantes idades em que se admite poder também ocorrer o mesmo tipo de transição. Visto que pretendemos assegurar a possibilidade das transições ocorrerem ao longo de mais de um intervalo etário quinquenal, podendo abranger até dois grupos etários, e nos interessa integrar nas projeções o estudo da tendência destas transições ocorridas em idades distintas, optamos pela segunda abordagem, seguida também por Goujon et al. (2007).

³¹Como discutido na secção 5.2.3.4, admite-se que a conclusão do Ensino Básico ocorra, na maioria dos casos, por volta dos 15 anos.

Consideremos, então, a variável resposta Y que designa o estado que cada elemento da população pode ocupar, em particular, o seu nível de escolaridade. Dada a hierarquia existente entre as várias categorias de escolaridade, Y é uma variável medida em escala ordinal; para além disso, cada indivíduo que atinja determinado nível de escolaridade, tem de ter passado, necessariamente, pelos níveis mais baixos. O que pretendemos é modelar as proporções de indivíduos em cada nível de escolaridade em função do tempo, de forma a projetar depois, para o futuro, os valores destas proporções de estado³².

Existem várias opções no que toca à forma de modelar o comportamento de uma variável resposta medida em escala ordinal. O modelo de *odds* proporcionais (*proportional odds model*), também conhecido como modelo cumulativo (*cumulative logit model*), é o mais frequentemente utilizado, inclusivamente em áreas como a epidemiologia ou a sociologia (Abreu et al., 2008); este modelo compara a probabilidade de uma resposta Y ser igual ou menor que uma determinada categoria com a probabilidade da resposta ser maior que esta categoria (Agresti, 2002). Contudo, no caso particular em que a variável resposta é determinada por um mecanismo sequencial, com a passagem obrigatória de cada indivíduo que chega a um nível pelos níveis inferiores, o modelo de razão contínua (*continuation-ratio logit model*) mostra-se mais adequado. O modelo de razão contínua relaciona o *odds* de estar num certo estado com o *odds* de estar nessa categoria ou acima. O modelo baseia-se em probabilidades condicionais: estima a probabilidade de estar numa categoria, dado que o indivíduo se encontra nessa categoria ou acima; para além disso, permite estimar a probabilidade condicional de estar acima de uma dada categoria dado que o indivíduo atingiu essa categoria específica, visto que estas duas probabilidades são complementares (Liu, 2010).

Dado que a variável Y tem $k = 4$ categorias, podemos definir as seguintes probabilidades condicionais:

$$\omega_j = P(Y = j | Y \geq j) = \frac{P(Y = j)}{P(Y = j) + \dots + P(Y = 4)} \quad j = 1, 2, 3$$

O modelo de razão contínua para Y pode ser formulado, considerando somente uma única covariável, como:

$$\text{logit}[\omega_j] = \log\left(\frac{\omega_j}{1 - \omega_j}\right) = \log\left(\frac{P(Y = j)}{P(Y > j)}\right) = \alpha + \beta X \quad j = 1, 2, 3$$

Pretendemos modelar as probabilidades ω_j em função do tempo, sendo a covariável X designada antes por T (tempo), por representar, por cada unidade, o avanço de um ano no período em análise. Na i -ésima realização de T , t_i , seja $\{y_{ij}, j = 1, \dots, 4\}$ a contagem de indivíduos em cada estado nesse ano, com o total de indivíduos dado por $n_i = \sum_j y_{ij}$ (Agresti, 2002). Se $b(n, y; \omega)$ é a probabilidade binomial de y sucessos em n provas de parâmetro ω , e se expressarmos a probabilidade multinomial

³²A modelação da tendência para as proporções de estado baseia-se nos dados do Inquérito ao Emprego (INE, I.P.), relativos ao número de indivíduos com mais de 15 anos, por sexo, grupos etários quinquenais e nível de escolaridade, para os anos entre 1998 e 2010. Dado que se pretende projetar esta tendência até ao ano de 2031, será de esperar que o erro de previsão seja muito elevado, dado o reduzido número de observações utilizadas para a modelação, relativamente ao número de períodos à frente para a previsão.

de (y_{i1}, \dots, y_{i4}) na forma $p(y_{i1}) p(y_{i2}|y_{i1}) p(y_{i3}|y_{i1}, y_{i2}) p(y_{i4}|y_{i1}, y_{i2}, y_{i3})$, a função massa de probabilidade da multinomial pode ser escrita como:

$$P(Y_{i1} = y_{i1}, \dots, Y_{i4} = y_{i4}) = b[n_i, y_{i1}; \omega_1(t_i)] b[n_i - y_{i1}, y_{i2}; \omega_2(t_i)] b[n_i - y_{i1} - y_{i2}, y_{i3}; \omega_3(t_i)]$$

A verosimilhança completa de n é, então, o produto das funções massa multinomiais para os diferentes valores de T . Quando as três regressões têm parâmetros distintos, é possível ajustar separadamente cada modelo, obtendo-se os mesmos resultados que no caso de se proceder à maximização conjunta da verosimilhança. Dado que cada um dos *logits* se refere a uma resposta binária, pode ser utilizado, por exemplo, o modelo de regressão logística binária em cada caso (Agresti, 2002).

Após estimação dos parâmetros dos três modelos, podemos estimar os valores das probabilidades de pertencer a cada um dos estados nos anos futuros. Com base nesses valores, calculam-se as proporções de transição correspondentes (aplicando a metodologia abordada na secção 5.2.3.4). Para estimar $p_{2,3,t}(10, 15)$ foi ainda necessário modelar a tendência relativa à taxa real de escolarização no Ensino Secundário e à taxa de retenção e desistência no mesmo nível de ensino, com base nas estimativas disponíveis (valores anuais entre 1998 e 2010), sendo depois utilizada a média dos valores projetados para os anos relativos a cada quinquénio de projeção. No primeiro caso, o melhor ajustamento (medido com base no coeficiente de determinação, R^2) foi obtido para uma linha de tendência logarítmica ($R^2 = 0,176$) e, para a taxa de retenção e desistência, o melhor ajustamento verificou-se com uma linha de tendência log-linear ($R^2 = 0,798$).

A construção destes modelos foi a forma encontrada de analisar a evolução das proporções de transição ao longo do tempo, de modo a projectá-la para o futuro, ainda que estejam sujeitos à escassez de informação de base.

Os resultados obtidos para as proporções de transição são apresentados na tabela 5.3.5, na qual se indicam também as proporções referentes ao quinquénio 2001-2006, utilizadas ao longo de todo o horizonte de projeção no cenário constante.

TABELA 5.3.5. *Proporções de transição entre níveis de escolaridade, por sexo, 2001-2031*

Fonte: Cálculos do autor

Quinquénio	H							M						
	$\hat{p}_{1,2}(5, 10)$	$\hat{p}_{1,2}(10, 15)$	$\hat{p}_{1,3}(10, 15)$	$\hat{p}_{2,3}(10, 15)$	$\hat{p}_{2,3}(15, 20)$	$\hat{p}_{3,4}(15, 20)$	$\hat{p}_{3,4}(20, 25)$	$\hat{p}_{1,2}(5, 10)$	$\hat{p}_{1,2}(10, 15)$	$\hat{p}_{1,3}(10, 15)$	$\hat{p}_{2,3}(10, 15)$	$\hat{p}_{2,3}(15, 20)$	$\hat{p}_{3,4}(15, 20)$	$\hat{p}_{3,4}(20, 25)$
2001-2006	0,016	0,526	0,089	0,386	0,672	0,395	0,414	0,016	0,595	0,142	0,386	0,815	0,744	0,432
2006-2011	0,016	0,568	0,129	0,493	0,811	0,719	0,416	0,016	0,625	0,184	0,493	0,832	0,987	0,536
2011-2016	0,016	0,577	0,173	0,531	0,786	0,691	0,306	0,016	0,634	0,221	0,531	0,773	0,933	0,540
2016-2021	0,016	0,571	0,225	0,564	0,782	0,698	0,250	0,016	0,631	0,259	0,564	0,737	0,935	0,553
2021-2026	0,016	0,551	0,285	0,590	0,769	0,702	0,184	0,016	0,618	0,300	0,590	0,701	0,937	0,556
2026-2031	0,016	0,517	0,352	0,609	0,750	0,707	0,105	0,016	0,597	0,342	0,609	0,665	0,938	0,549

Destaca-se, na análise destes valores, o facto de a proporção de indivíduos que transita do Ensino Básico para o Ensino Secundário no grupo etário entre os 15 e os 20 anos, $p_{2,3,t}(15, 20)$, diminuir

ao longo do tempo, para ambos os sexos; no entanto, esta descida é compensada pelo aumento das transições entre estes níveis de ensino nos grupos etários mais jovens. Interessa ainda relembrar que as proporções de transição para o Ensino Básico nos indivíduos mais jovens se mantêm constantes devido ao facto de dispormos somente da proporção de pessoas entre os 10 e os 14 anos nesse nível de ensino com respeito ao ano de 2001.

5.4. Software utilizado na elaboração das projeções

A implementação da metodologia formulada para as projeções teve como suporte o *Microsoft Office Excel* 2010, que foi também utilizado para construir os gráficos e tabelas apresentados na secção seguinte. O *software* R (versão 2.13.1) foi utilizado na construção das tabelas de mortalidade associadas ao cenário central do INE, I.P., através da biblioteca *LifeTables*, e na modelação da tendência relativa aos níveis de escolarização da população portuguesa (biblioteca VGAM).

Resultados

Neste capítulo, são apresentados os resultados obtidos nas projeções, no que diz respeito à dimensão e estrutura da população residente em Portugal até ao ano de 2031. Dado que os dois cenários considerados se distinguem nas estimativas das proporções de transição entre níveis de escolaridade, é na distribuição por nível de escolaridade, tanto em cada coorte como no conjunto da população, que se manifestam as principais diferenças nos resultados. As ligeiras diferenças verificadas entre os cenários no que se refere à dimensão da população total e de cada grupo por sexo e idade, são fruto dos diferenciais de fecundidade, mortalidade e migrações, por grau de instrução.

6.1. População residente por sexo, grupo etário e nível de escolaridade, 2001-2031

Em ambos os cenários é esperado um aumento do número total de residentes em Portugal, passando dos 10.256.658 habitantes verificados em 2001, para 10.737.578 de acordo com o cenário constante ou para 10.741.559 habitantes segundo o cenário de tendência, em 2031. A distribuição dos residentes por sexo deverá manter-se estável ao longo do tempo: em 2001, 51,7% dos residentes eram do sexo feminino e, em 2031, esta percentagem deverá ser de aproximadamente 51,3%, seja qual for o cenário considerado. Já no que diz respeito à representatividade de cada grupo etário no total da população, esperam-se diferenças consideráveis em relação ao ano de partida: a proporção de indivíduos com mais de 65 anos deverá aumentar até cerca de 21,3%, enquanto a proporção de crianças e jovens com menos de 14 anos deverá situar-se ligeiramente acima dos 14%, em 2031 (figura 6.1.1).

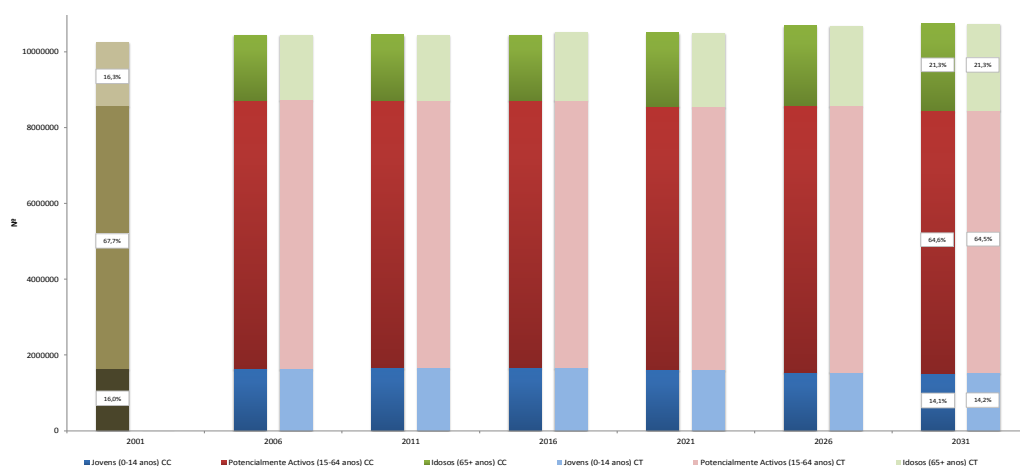


FIGURA 6.1.1. Grupos funcionais, 2001-2031, por cenário

CC: Cenário Constante, CT: Cenário de Tendência

Fonte: Cálculos do autor

Como consequência do aumento da proporção de idosos relativamente à percentagem de jovens na população, o índice de envelhecimento deverá aumentar consideravelmente. De acordo com o cenário de tendência, em 2031, existirão 150,3 pessoas com mais de 65 anos por cada 100 indivíduos abaixo dos 15 anos de idade (figura 6.1.2). A população ativa tornar-se-á também cada vez mais envelhecida

e, em 2031, existirão apenas 75 pessoas entre os 20 e os 29 anos por cada 100 pessoas com idade entre os 55 e os 64 anos. Já no grupo das mulheres em idade reprodutiva, deverá existir uma ligeira recuperação no que respeita à representatividade das mulheres mais jovens (entre os 15 e os 34 anos) relativamente às mais velhas (entre os 35 e os 49 anos), não chegando, contudo, a existir um maior número de mulheres do primeiro grupo em relação ao segundo.

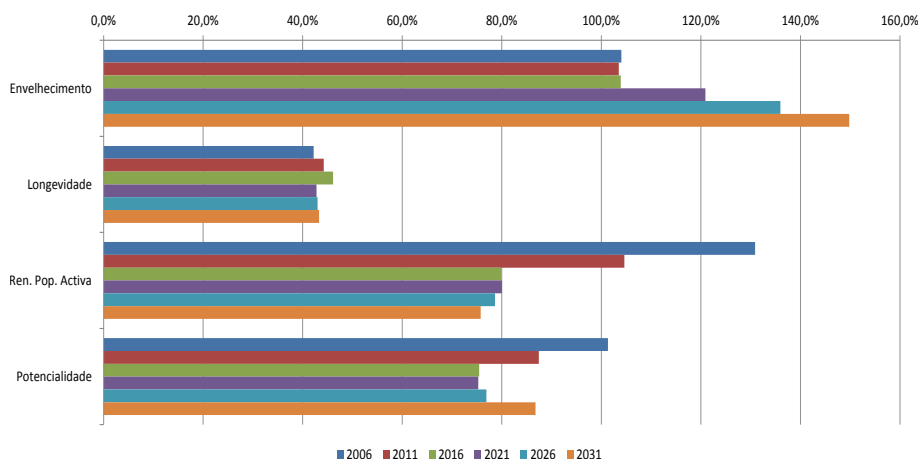


FIGURA 6.1.2. Índices-resumo, 2006-2031, Cenário Tendência

Fonte: Cálculos do autor

A proporção de residentes que completam pelo menos o Ensino Básico deverá continuar a aumentar, sendo possível esperar que, em 2031, mais de metade da população tenha concluído este nível de ensino (figura 6.1.3). Mesmo com a manutenção das proporções de transição (cenário constante), a escolarização da população tenderá a aumentar ao longo do tempo, à medida que as coortes mais jovens, e mais escolarizadas, substituem as coortes mais velhas. O cenário de tendência resulta numa aceleração deste aumento, por se basear em proporções de transição que seguem a tendência da última década e, como tal, crescem ao longo do período de projeção.

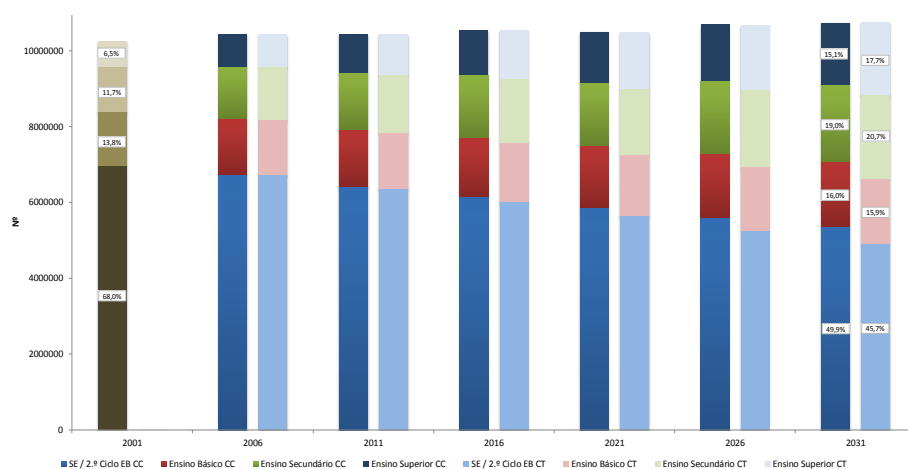


FIGURA 6.1.3. Número de residentes por nível de escolaridade, 2001-2031, por cenário

CC: Cenário Constante, CT: Cenário de Tendência

Fonte: Cálculos do autor

Dado que a conclusão do Ensino Básico ocorre habitualmente em torno dos 15 anos de idade, o grupo das crianças e jovens até aos 14 anos permanece, na sua grande maioria, no primeiro grupo, independentemente da passagem do tempo e dos níveis de escolarização considerados. Desta forma, importa analisar os efeitos da escolarização focando-nos nos indivíduos com idades dos 15 anos em diante. Considerando apenas os residentes no grupo etário dos 15 aos 64 anos (figura 6.1.4), verifica-se que, apesar de o número de efetivos nestas idades se manter estável, a valorização do capital humano neste grupo deverá ser evidente, com a queda da proporção de indivíduos que não completaram o Ensino Básico de 55,5%, em 2001, para aproximadamente 30%, em 2031 (34,5% segundo o cenário constante e 28,1% no cenário de tendência). Já a percentagem de indivíduos que possui um curso de Ensino Superior, naquele que é tido como o grupo etário correspondente à idade ativa, deverá atingir valores próximos dos 20%, em 2031, quando 30 anos antes era de apenas 9%.

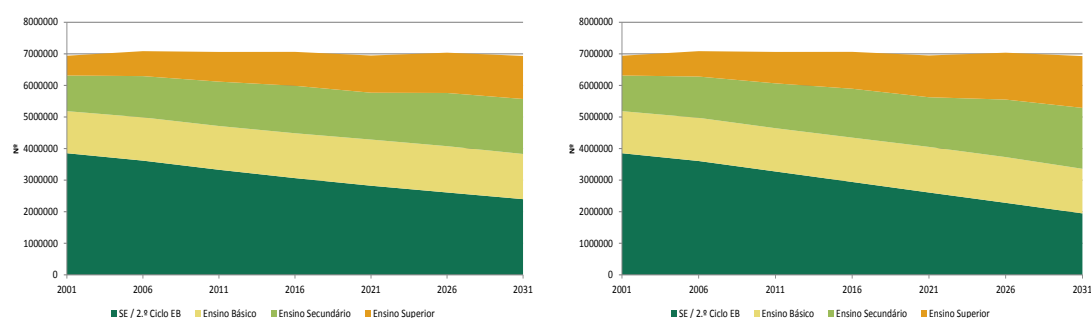


FIGURA 6.1.4. Número de residentes com 15-64 anos, por nível de escolaridade, 2001-2031, por cenário

Cenário Constante (esquerda) e Cenário Tendência (direita)

Fonte: Cálculos do autor

Na figura 6.1.5, baseada no cenário de tendência, verifica-se o aumento do número total de pessoas com 65 ou mais anos, acompanhado pela clara alteração da sua estrutura em termos educacionais: em 2031, 30,6% dos indivíduos deste grupo deverão ter completado pelo menos o 3.º Ciclo do Ensino Básico, o que representa uma subida de 25,7 pontos percentuais relativamente a 2001; já a percentagem de pessoas com um curso de Ensino Superior completo, que se situava nos 2,4% em 2001, deverá aumentar até 11,3% em 2031.

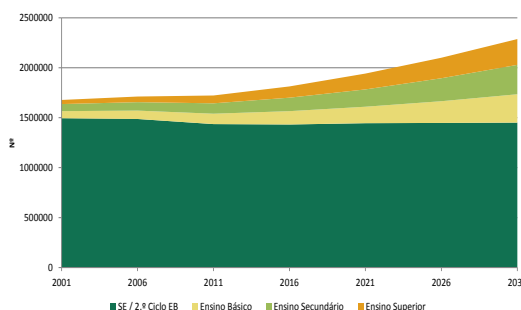


FIGURA 6.1.5. Número de residentes com 65 ou mais anos, por nível de escolaridade, 2001-2031, Cenário Tendência

Fonte: Cálculos do autor

A distribuição dos residentes pelo conjunto das três variáveis - sexo, grupo etário e nível de escolaridade - pode ser analisada através da construção de pirâmides etárias multiestado. A comparação entre a estrutura populacional observada em 2001 e a estimada para o ano de 2031 (figuras 6.1.6 e 6.1.7) indica que a proporção de indivíduos que completam níveis de escolaridade mais elevados deverá aumentar em todos os grupos etários de ambos os sexos, segundo qualquer um dos cenários. As pirâmides relativas ao final de cada quinquênio, para os dois cenários, são apresentadas no apêndice B.

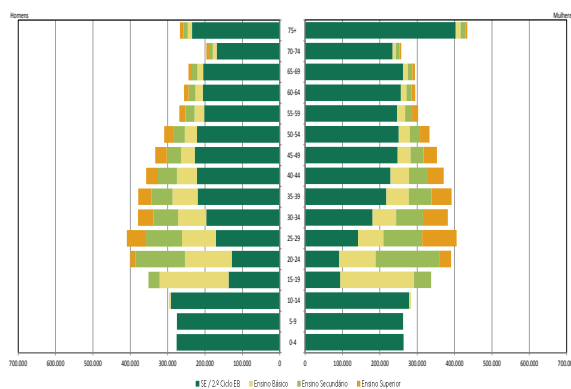


FIGURA 6.1.6. *População portuguesa por grupo etário, sexo e nível de escolaridade, 1 de janeiro de 2001*

Fonte: INE, Estimativas Anuais da População Residente e cálculos do autor

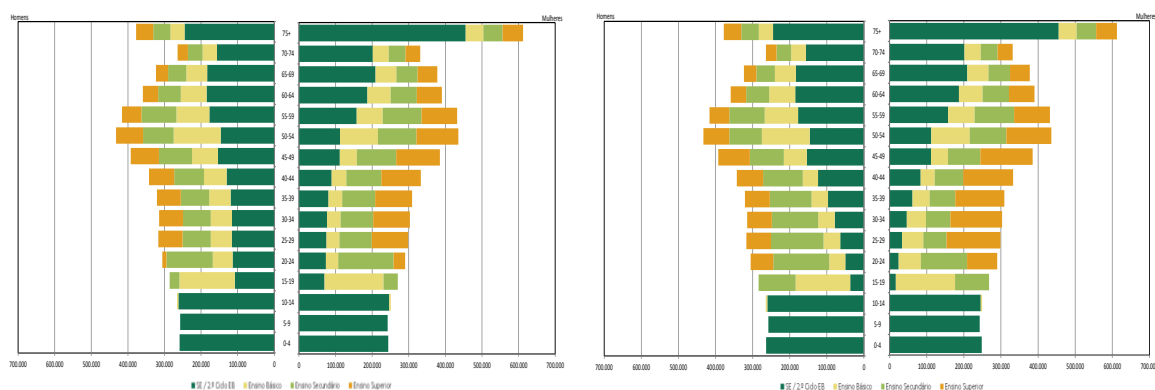


FIGURA 6.1.7. *População portuguesa por grupo etário, sexo e nível de escolaridade, 2031*

Cenário Constante (esquerda) e Cenário Tendência (direita)

Fonte: Cálculos do autor

Em qualquer um dos cenários, permanecem diferenças entre indivíduos de sexos distintos, no que respeita ao seu padrão de escolarização. Apesar de nas coortes mais velhas (acima dos 70 anos) se notar ainda, em 2031, uma maior percentagem de homens com pelo menos o Ensino Básico, relativamente às mulheres, a escolarização entre as mulheres tende a superar os níveis de escolarização masculina, o que é evidente nas coortes mais jovens. Considerando todos os grupos etários, a percentagem de homens que não chegaram a completar o Ensino Básico era, em 2001, ligeiramente inferior relativamente às mulheres (67,7% dos homens e 68,3% das mulheres), situação que se modifica ao longo do período de projecção, com 47,4% de homens que não têm o Ensino Básico completo, em 2031, contra 44,2% das mulheres (cenário de tendência); entre as mulheres, apenas 7,3% tinha concluído algum grau de

Ensino Superior em 2001, valor que deverá subir para aproximadamente 22% em 2031; já entre os homens, este valor era de apenas 5,7% em 2001 e deverá também subir até 2031, atingindo os 13,2% nesse ano. Dado que a passagem a um nível de escolaridade mais elevado é um evento irreversível, podemos esperar que a proporção de idosos com níveis de instrução superiores continue a aumentar, já que a escolarização das coortes mais jovens parece continuar a evoluir positivamente (a proporção de pessoas com idades entre os 30 e os 34 anos com o Ensino Superior completo, em 2031, deverá aproximar-se do dobro da percentagem verificada em 2001).

CAPÍTULO 7

Discussão

Regressando ao esquema proposto por Willekens (1990) para a elaboração estruturada de projeções demográficas, resta-nos, nesta fase, proceder à avaliação dos resultados das projeções, com base nos dados observados nos anos já decorridos. Este tipo de análise é possível por termos iniciado o período de projeção em 2001, reservando assim um intervalo de tempo considerável para o qual já dispomos de dados empíricos; caso contrário, a avaliação das projeções só poderia ser feita num momento futuro. Proceder-se-á então à comparação entre os resultados obtidos nestas projeções multiestado, os resultados relativos ao cenário central das projeções elaboradas pelo INE, I.P., e os dados provisórios do XV Recenseamento Geral da População (Censos 2011).

Para além disso, serão discutidas neste capítulo as principais limitações associadas à elaboração do trabalho apresentado, bem como a sua relevância e as intenções relativas a trabalhos futuros.

7.1. Avaliação de resultados

O interesse primordial das projeções elaboradas foi o de obter estimativas da dimensão e estrutura da população residente em Portugal por sexo, grupo etário e nível de escolaridade para os anos futuros. A análise comparativa dos resultados obtidos em relação a outras projeções ou a dados reais é limitada, por um lado devido à inexistência de outras projeções que considerem as três variáveis em simultâneo e, por outro, porque não se encontram ainda disponíveis os resultados definitivos dos Censos 2011 com o mesmo nível de desagregação. Ainda assim, é possível efetuar algumas comparações, tendo 2011 como ano de referência, dado ser aquele relativamente ao qual se encontram disponíveis as últimas estimativas divulgadas.

No que se refere ao número total de residentes, o cenário de tendência das projeções multiestado aponta para um valor inferior ao verificado nos Censos 2011 (10.445.127 e 10.561.614 habitantes, respetivamente), ao contrário do cenário central das projeções do INE, I.P., segundo o qual a população residente em 2011 é de 10.675.295 indivíduos. Apesar destas diferenças, o erro associado às duas projeções, em relação aos dados obtidos no recenseamento, é equivalente em termos proporcionais, correspondendo a 1,1% da população observada nos Censos.

Distribuindo a população por sexo e grupo etário (tabela 7.1.1), verificamos que o cenário de tendência das projeções apresentadas resulta em valores relativos ao número de crianças e jovens até aos 14 anos que se situam acima dos resultados provisórios dos Censos 2011 e dos resultados das projeções do INE,

I.P., para ambos os sexos. Por outro lado, os números projetados de indivíduos com 65 ou mais anos situam-se abaixo dos observados no recenseamento e dos projetados pelo INE, I.P.. Já em relação às pessoas com idades entre os 15 e os 64 anos, os resultados das projeções multiestado aproximam-se mais dos dados dos Censos do que as projeções do INE.

TABELA 7.1.1. *População residente (N.º), por sexo e grupo etário, 2011, comparação de resultados (Censos 2011, Cenário central INE e Cenário tendência)*

Fonte: INE, XV Recenseamento Geral da População (resultados provisórios); INE, projeções de População Residente (Portugal, 2008-2060); Cálculos do autor

Datas de referência: Censos 2011 - 21/03/2011; Cenário central INE e Cenário tendência - 01/01/2011

Grupo etário	H			M			HM		
	Censos 2011	Cenário central INE	Cenário tendência	Censos 2011	Cenário central INE	Cenário tendência	Censos 2011	Cenário central INE	Cenário tendência
0-14	804133	831197 (+3,4%)	855497 (+6,4%)	768413	785434 (+2,2%)	808159 (+5,2%)	1572546	1616631 (+2,8%)	1663656 (+5,8%)
15-24	580834	589226 (+1,4%)	584887 (+0,7%)	564936	572004 (+1,3%)	560983 (-0,7%)	1145770	1161230 (+1,3%)	1145869 (+0,0%)
25-64	2813232	2931085 (+4,2%)	2908758 (+3,4%)	3007562	3031893 (+0,8%)	3004796 (-0,1%)	5820794	5962978 (+2,4%)	5913553 (+1,6%)
65+	849188	805502 (-5,1%)	705239 (-17,0%)	1173316	1128954 (-3,8%)	1016810 (-13,3%)	2022504	1934456 (-4,4%)	1722049 (-14,9%)

O número de indivíduos em cada nível de escolaridade obtido no cenário de tendência difere dos resultados provisórios dos Censos 2011 sobretudo nos grupos correspondentes ao Ensino Básico e ao Ensino Superior, sendo, em relação a qualquer um destes grupos e a ambos os sexos, inferior nas projeções elaboradas (tabela 7.1.2).

TABELA 7.1.2. *População residente (N.º), por sexo e nível de escolaridade mais elevado completo, 2011, comparação de resultados (Censos 2011 e Cenário tendência)*

Fonte: INE, XV Recenseamento Geral da População (resultados provisórios); Cálculos do autor

Datas de referência: Censos 2011 - 21/03/2011; Cenário tendência - 01/01/2011

Nível de escolaridade	H		M		HM	
	Censos 2011	Cenário tendência	Censos 2011	Cenário tendência	Censos 2011	Cenário tendência
SE / 2.º Ciclo EB	2934714	3096840 (+5,5%)	3171962	3267732 (+3,0%)	6106676	6364572 (+4,2%)
Ensino Básico	882766	783767 (-11,2%)	804319	699823 (-13,0%)	1687085	1483590 (-12,1%)
Ensino Secundário	731927	746232 (+2,0%)	773477	776950 (+0,4%)	1505404	1523182 (+1,2%)
Ensino Superior	497980	427542 (-14,1%)	764469	646242 (-15,5%)	1262449	1073784 (-14,9%)

No que diz respeito a dados sobre a população desagregada por sexo, grupo etário e nível de escolaridade, dispomos somente das estimativas anuais relativas ao Inquérito ao Emprego elaborado pelo INE, I.P.. Em relação à população com 15 ou mais anos no ano de 2011 (tabela 7.1.3), as projeções multiestado (cenário de tendência) apontam para números de residentes com o Ensino Básico e com o Ensino Superior inferiores aos estimados pelo INE, sendo em relação a estes grupos de escolaridade que se observam as maiores diferenças. O número de residentes por grupo etário difere principalmente no que toca às pessoas com 65 ou mais anos, que é, para cada um dos sexos, mais baixo segundo os resultados das projeções. As projeções resultam, regra geral, em valores mais elevados de pessoas que não concluíram o Ensino Básico e em valores mais baixos de indivíduos nos restantes grupos, para o mesmo grupo etário e independentemente do sexo considerado. Tal não acontece unicamente no grupo

das pessoas com 65 ou mais anos, em que, nas projeções, o número de indivíduos sem o Ensino Básico é muito inferior e o daqueles que completaram o Ensino Secundário é superior ao estimado pelo INE, I.P..

TABELA 7.1.3. *População residente (N.^o), por sexo, grupo etário e nível de escolaridade, 2011, comparação de resultados (Inquérito ao Emprego e Cenário tendência)*

Fonte: INE, Inquérito ao Emprego (Série 2011); Cálculos do autor

Grupo etário	H					M					HM					
	SE / 2.ºCEB	E. Bás.	E. Sec.	E. Sup.	TOTAL	SE / 2.ºCEB	E. Bás.	E. Sec.	E. Sup.	TOTAL	SE / 2.ºCEB	E. Bás.	E. Sec.	E. Sup.	TOTAL	
Inq. ao Emprego	15-24	100600	265000	192000	26600	584200	61900	230700	210700	55300	558600	162500	495700	402700	81900	1142800
	25-34	163500	226000	220100	167600	777200	108900	180600	221200	245200	755900	272400	406600	441300	412800	1533100
	35-44	341400	199400	146900	121600	809300	274200	177500	179300	179100	810100	615600	376900	326200	300700	1619400
	45-64	840000	223200	147200	136800	1347200	880000	253300	143300	177000	1453600	1720000	476500	290500	313800	2800800
	65+	669500	63000	30200	45700	808400	999100	53200	27400	52600	1132300	1668600	116200	57600	98300	1940700
	TOTAL	2115000	976600	736400	498300	4326300	2324100	895300	781900	709200	4710500	4439100	1871900	1518300	1207500	9036800
Cenário tendência	15-24	190893	191253	181434	21306	584887	119630	197479	198538	45335	560983	310523	388733	379972	66641	1145869
	25-34	276857	184829	174424	148975	785086	196529	140759	177602	252037	766926	473386	325588	352026	401012	1552012
	35-44	379585	168791	164360	95602	808339	337291	134417	179481	163742	814931	716876	303208	343841	259345	1623270
	45-64	842832	181512	171198	119791	1315333	930181	171997	172137	148624	1422938	1773013	353509	343335	268415	2738272
	65+	555508	53049	54815	41867	705239	880234	50879	49192	36504	1016810	1435741	103928	104008	78372	1722049
	TOTAL	2245676	779434	746232	427542	4198884	2463864	695532	776950	646242	4582588	4709540	1474966	1523182	1073784	8781472

A utilização dos diferenciais de mortalidade infantil por escolaridade da mãe para estimar as diferenças nos níveis de mortalidade entre os grupos de escolaridade, bem como as aproximações relativas à estimação do saldo migratório, deverão contribuir para explicar as disparidades observadas no que toca à distribuição dos indivíduos mais velhos pelos quatro grupos, dado que, como os indivíduos com 65 ou mais anos em 2011 tinham pelo menos 45 anos no início do período de projeção, não ocorreram, neste grupo, transições entre níveis de escolaridade durante a década 2001-2011. Por outro lado, no grupo etário dos 15-24 anos, o cenário tendência indica uma proporção de indivíduos que não completaram o Ensino Básico muito superior à estimada pelo Inquérito ao Emprego, enquanto o número de indivíduos com Ensino Básico se situa muito abaixo dos resultados desse estudo. Neste caso, será uma subestimação das probabilidades de transição para este nível de ensino (em parte devido à utilização das probabilidades estimadas para o grupo etário dos 10-14 anos, em que foi necessário recorrer unicamente a dados dos Censos 2001) que melhor explica estas diferenças. Apesar do efeito das mesmas sobre o número total de pessoas entre os 15 e os 24 anos ser diminuto (pois os níveis de mortalidade dos vários grupos de escolaridade são muito baixos e ainda muito próximos nestas idades), podem ajudar a explicar o elevado número de jovens (com idades até aos 14 anos), por serem as mulheres com menores níveis de escolaridade as que mais contribuem para a natalidade, entre os 15 e os 24 anos.

No que respeita aos resultados obtidos, interessa ainda referir que o cenário tendência sugere o cumprimento da meta estabelecida no âmbito do programa *Metas Educativas 2021: A Educação que queremos para a geração dos Bicentenários* (ver secção 4.3), segundo a qual se pretende que, até 2021,

entre 60% a 90% de jovens terminem o Ensino Secundário: de acordo com este cenário, a percentagem de indivíduos entre os 25 e os 29 anos que completaram pelo menos o Ensino Secundário situar-se-á, em 2021, em 62,3%. Pelo contrário, o objetivo estabelecido pelo programa *Educação e Formação para 2020*, de que a percentagem de adultos de 30-34 anos com nível de ensino superior deva ser de pelo menos 40% em 2020, poderá não ser cumprido: em 2021, com base no cenário tendência, esta percentagem atingirá somente os 31,3%.

7.1.1. Como quantificar a incerteza dos resultados? Como discutido na secção 2.2.1, a quantificação da incerteza na elaboração de projeções só pode resultar da aplicação de projeções probabilísticas. A metodologia seguida neste trabalho baseou-se numa abordagem determinística que é, de resto, a opção tomada pela grande maioria dos demógrafos, tanto em relação às projeções por idade e sexo como às projeções multidimensionais. Entre os dois cenários construídos, admitimos que o cenário de tendência seja o mais plausível, pois parece irrealista pensar que o nível de escolarização se mantenha até 2031 em níveis semelhantes aos verificados no início dos anos 2000. Mas de que forma poderíamos avaliar esta plausibilidade? Ao contrário da projeção determinística, que apenas prevê um número de residentes (por cada grupo) para um determinado ano, uma projeção probabilística pode indicar quão plausível é que o número de residentes em cada grupo se situe num determinado intervalo de valores. Alho e Keilman (2010) apresentou uma forma de obter uma previsão probabilística a partir de uma projeção determinística multiestado³³. Esta metodologia passa por calcular os erros associados a uma projeção anterior que seja semelhante, a partir da sua comparação com os resultados de recenseamentos, por exemplo, e modelar a distribuição destes erros, assumindo posteriormente que a mesma se aplica à projeção multiestado atual. No nosso caso concreto, considera-se que, apesar da inexistência de uma projeção multiestado idêntica, seria possível especificar as distribuições preditivas das proporções de indivíduos em cada grupo de escolaridade com base na comparação entre os resultados obtidos e os dados dos Censos 2011, que constituiriam a informação empírica através da qual se procederia à simulação dos erros de previsão para os restantes anos de projeção. À semelhança do trabalho de Alho e Keilman (2010), os valores esperados das proporções seriam obtidos a partir do cenário de tendência da projeção determinística, considerando-se as proporções correspondentes como estimativas pontuais dos mesmos. Por fim, seria elaborada uma previsão do número de pessoas num dado nível de escolaridade para cada final de quinquénio como o produto de duas variáveis aleatórias, nomeadamente a proporção prevista para esse nível de escolaridade e o total da população previsto por uma projeção estocástica. Para tal, poderiam ser utilizados os resultados da projeção estocástica da população portuguesa elaborada no âmbito do projeto *Uncertain population of Europe*, que fornece distribuições preditivas da população por sexo e idade para 18 países europeus, para o período 2004-2050. Alho (2006) apresenta uma descrição geral dos resultados destas projeções, sendo os pressupostos discutidos por Alders et al. (2007) e Alho et al. (2008). Os resultados das simulações

³³O trabalho apresentado por Alho e Keilman (2010) refere-se a projeções da população norueguesa por sexo, grupo etário e tipo de agregado familiar; no entanto, os autores sugerem que um método similar, com as adaptações necessárias, pode ser utilizado para obter projeções probabilísticas para outros tipos de divisão da população.

efetuadas, que se encontram disponíveis em www.stat.fi/tup/euue/no11-results-nor.html, seriam então combinados com simulações realizadas sobre a distribuição preditiva de cada proporção. Outra opção seria utilizar os resultados das projeções probabilísticas realizadas por Pedrosa (2011). No entanto, dado que os resultados dos Censos 2011 relativos à população residente por sexo, grupo etário e nível de escolaridade ainda não se encontram disponíveis, não foi possível a aplicação desta metodologia em tempo útil.

7.2. Limitações

Podem ser apontadas algumas limitações relativas ao trabalho apresentado, associadas sobretudo à complexidade inerente à elaboração de projeções demográficas.

O primeiro ponto a discutir deverá ser a incerteza subjacente aos resultados obtidos. Neste trabalho, optou-se pela elaboração de projeções determinísticas, dado ser esta a abordagem mais frequente na elaboração de projeções, e por se considerar que não se dispunham dos dados suficientes para a construção de projeções probabilísticas. Como referido por O'Neill et al. (2001) e discutido na secção 2.2.1, a elaboração de projeções probabilísticas exige a definição de distribuições de incerteza para cada uma das componentes demográficas e, no caso das projeções multiestado, também para as transições entre estados; as técnicas geralmente utilizadas para estimar estes níveis de incerteza consistem na análise de séries temporais, na análise dos erros de projeções anteriores ou na avaliação de argumentos de especialistas resultando em conjuntos de distribuições probabilísticas subjetivas. A inexistência de séries temporais relativas à fecundidade por nível de escolaridade suficientemente longas e a ausência de dados sobre a mortalidade e migrações por grau de instrução colocam de lado esta possibilidade. O recurso à segunda opção pode ser uma alternativa viável quando se encontrarem disponíveis os dados definitivos dos Censos 2011 por sexo, grupos etários quinquenais e nível de escolaridade, utilizando-se, por exemplo, a metodologia proposta por Alho e Keilman (2010), tal como discutido na secção anterior. Por fim, a falta de uma metodologia consensual com vista à avaliação dos argumentos de especialistas e as limitações de tempo associadas à elaboração desta tese impossibilitaram a aplicação da terceira alternativa.

A interpretação dos resultados exige que se tenha em mente a natureza condicional dos mesmos, por terem por base cenários do tipo “se...então”, característica associada a todas as projeções determinísticas. Como acontece na generalidade das projeções, assume-se que não ocorrerá, no futuro, nenhuma alteração profunda nas circunstâncias que determinam os comportamentos demográficos e, no nosso caso concreto, de escolarização (como, por exemplo, guerras ou catástrofes naturais). Para além disso, é difícil prever que medidas governativas poderão vir a ter impacto sobre o grau de escolarização da população, dado que este tipo de efeitos surge normalmente no médio a longo prazo, ou de que forma o comportamento migratório (geral e por nível de instrução) se alterará, em particular levando em consideração o contexto socioeconómico atual. Ainda que o nível de incerteza associado às projeções

não possa ser medido, é um facto que o avanço no horizonte temporal de projecção conduz ao aumento desta incerteza, pelo que se exige prudência na leitura e interpretação de resultados a longo prazo.

A exigência de dados para a aplicação de modelos demográficos multiestado não pôde ser cumprida integralmente. Como exemplo, a inexistência de registos relativos ao grau de instrução dos indivíduos falecidos em Portugal ou dos seus migrantes impossibilitou a estimação direta dos diferenciais de mortalidade ou migrações por nível de escolaridade. Procurou-se, tanto quanto possível, contornar a escassez de informação de base, através de técnicas utilizadas em trabalhos semelhantes (Lutz et al., 2005; Goujon et al., 2007), por se julgar que as aproximações efetuadas conduzem, ainda assim, a melhores resultados do que os que obteríamos se ignorássemos estas vertentes da modelação multiestado. Da mesma forma, a estimação das probabilidades de transição entre níveis de escolaridade foi também dificultada pelo tipo de dados disponíveis, o que é comum em projecções deste tipo (Lutz e Goujon, 2001; Lutz et al., 2005; Goujon et al., 2007; Wils, 2007; KC et al., 2010).

Outros pontos merecem destaque: apesar do enfoque que a educação e formação na idade adulta têm vindo a receber nos últimos anos, nomeadamente como alvo essencial no âmbito da Estratégia de Lisboa, optou-se por não considerar a possibilidade dos indivíduos com idades superiores a 30 anos poderem passar para um nível de escolaridade mais elevado; assume-se que, na grande maioria dos casos em que existe formação na idade adulta, esta não é necessariamente formal e não conduz à passagem para um novo estado, de entre os níveis considerados (Goujon, 2008). De facto, a opção de limitar os graus de instrução aos quatro considerados prende-se com a necessidade de minimizar a complexidade do modelo e, em consequência, dos dados necessários à sua aplicação. Para além disso, é necessário ter em mente que a análise da população por nível de escolaridade mais elevado completo não reflete a qualidade da educação recebida, mas medições do capital humano nesse sentido são ainda muito escassas.

7.3. Relevância do trabalho e desenvolvimentos futuros

As projecções apresentadas constituem o primeiro exercício com o intuito de estimar a evolução, no longo prazo, da dimensão e estrutura da população residente em Portugal por sexo, grupo etário e nível de escolaridade, realizado ao nível nacional. A decisão de recorrer ao modelo de projecção multiestado baseou-se na intenção de captar, dentro do possível, o dinamismo próprio do comportamento populacional, em particular no que respeita à conclusão de graus de instrução mais elevados. Por outro lado, a utilização deste tipo de modelo permitiu também considerar as diferenças no comportamento dos indivíduos com níveis de escolaridade distintos, relativamente à fecundidade, mortalidade e migrações. Este tipo de projecções são pertinentes, não apenas para a tomada de decisão no contexto, por exemplo, do planeamento de políticas educacionais, mas também por permitirem o estudo mais detalhado de cada coorte definida pela idade e sexo, ao ser considerada adicionalmente uma das variáveis que mais contribuem para a heterogeneidade presente entre os indivíduos.

No contexto do Projeto de Investigação em que este trabalho se insere, pretende-se que as projeções elaboradas auxiliem a avaliação da forma como o aumento das despesas de saúde ligadas ao envelhecimento pode ser contrabalançado pelo efeito positivo da escolarização sobre o estado de saúde da população portuguesa. Para tal, pretender-se-á proceder à produção de projeções cuja unidade geográfica sejam as regiões NUTS II - Norte, Centro, Lisboa, Alentejo, Algarve, Região Autónoma dos Açores e Região Autónoma da Madeira - o que representará um desafio crescente, na medida em que será necessário serem consideradas, ainda, as eventuais migrações internas, entre as regiões. Estes fluxos migratórios poderão ser estimados com base nos dados recolhidos nos recenseamentos realizados pelo INE, I.P., relativos à residência na data do recenseamento e num momento anterior (com uma distância de cinco e dois anos em relação ao momento censitário). Com este tipo de informação retrospectiva, poderão ser estimadas matrizes de fluxos migratórios inter NUTS II, análogas às apresentadas por Magalhães (2003). A maior dificuldade consistirá, assim, na combinação simultânea das migrações internacionais, com as migrações internas entre as regiões e as transições entre níveis de escolaridade.

Numa fase posterior do Projeto, estudar-se-á a forma como as características demográficas e o nível de escolaridade se relacionam com o estado de saúde, recorrendo, para tal, à construção de modelos de regressão a partir dos dados do último Inquérito Nacional de Saúde (INE/INSA). Como resultado, espera-se conseguir avaliar o impacto das alterações nas estruturas demográficas por idade, sexo e grupo etário, por região, em particular no que respeita ao efeito que tais alterações poderão exercer sobre o estado de saúde da população. Os resultados deverão fornecer pistas sobre os ajustamentos necessários em termos de políticas sociais e de saúde, num tal contexto de mudança demográfica e social.

Conclusões

O estudo da população de um determinado país ou região é essencial como suporte para a tomada de decisão num conjunto de sectores que, por sua vez, influenciam direta ou indiretamente os comportamentos demográficos e, num prazo mais ou menos longo, a estrutura populacional futura. No que respeita, em particular, ao estado de saúde da população residente em Portugal, interessa perceber, desde já, de que modo o evidente envelhecimento demográfico e o aumento dos níveis de escolarização poderão contribuir para alterações futuras que exijam a adaptação ou formulação de políticas públicas, nomeadamente de âmbito social. O presente trabalho insere-se num Projeto que visa contribuir precisamente neste sentido. Foram elaboradas as primeiras projeções nacionais por sexo, grupo etário e nível de escolaridade, relativamente ao período 2001-2031, que servirão de base a um estudo completo das alterações esperadas sobre o estado de saúde, para o qual se recorrerá também a outras técnicas estatísticas, como a construção de modelos de regressão. Para a elaboração das referidas projeções foi utilizado um modelo de projeção multiestado, com vista a incorporar o modo dinâmico como se processa o comportamento demográfico. O modelo apresentado resultou de um conjunto de fatores: a) por um lado, do conhecimento adquirido relativamente aos modelos populacionais desenvolvidos até ao momento e aos tipos de projeções demográficas existentes; b) por outro, da análise retrospectiva da evolução demográfica da população portuguesa; ainda, c) da estrutura atual do Sistema Educativo Português e das reformas em curso e, por fim, d) do tipo de dados que se encontram disponíveis. Foram estabelecidos dois cenários de evolução - cenário constante e cenário de tendência - de forma a lidar com a incerteza associada aos resultados.

O cumprimento do principal objetivo deste trabalho foi marcado por um conjunto de limitações associadas à sua complexidade e à escassez de dados disponíveis, traduzindo-se em dificuldades comuns às observadas em trabalhos semelhantes; estes trabalhos, por sua vez, forneceram as pistas necessárias para que fosse possível contornar tais obstáculos e progredir de forma a atingir os resultados apresentados. A melhoria destes resultados poderá ser obtida num futuro próximo, através da utilização dos resultados definitivos dos Censos 2011 para uma avaliação cuidada e identificação das alterações necessárias ao método aplicado. Num prazo mais alargado, o desenvolvimento das técnicas de projeção multiestado e o registo de dados relativos aos falecimentos por nível de escolaridade, por exemplo, poderão contribuir também para aumentar a qualidade deste tipo de projeções.

A contribuição deste trabalho consiste sobretudo no fornecimento de dados que corroboram as expectativas de que a população portuguesa se venha a tornar cada vez mais envelhecida e, em simultâneo,

mais escolarizada, mas estende-se ainda à demonstração de como a integração de áreas como a Demografia, a Análise de Sobrevivência e a Análise de Regressão, traduzida na utilização do modelo multiestado e exigindo a utilização integrada de diferentes *softwares*, facilita o estudo de fenómenos tão complexos como o comportamento demográfico. Os modelos multiestado podem ser utilizados noutras aplicações, nomeadamente na modelação de fenómenos e de comportamentos que estejam associados à passagem entre diferentes estados (por exemplo, entre diferentes estadios de uma dada doença) e, como tal, o seu estudo e o desenvolvimento dos métodos de estimação, associados aos progressos em termos computacionais, poderá ser benéfica em diferentes campos.

Bibliografia

- Abreu, M. N. S. et al. (2008). Ordinal logistic regression models: application in quality of life studies. *Cadernos de Saúde Pública* 24, s581–s591. ISSN: 0102-311X. URL: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2008001600010&nrm=iso.
- Agresti, A. (2002). *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience. ISBN: 9780471360933.
- Ahuja, V. e D. Filmer (1995a). *Educational attainments in developing countries : new estimates and projections disaggregated by gender*. Policy Research Working Paper Series 1489. The World Bank. URL: <http://elibrary.worldbank.org/docserver/download/1489.pdf?expires=1341158460&id=id&accname=guest&checksum=161084D3E6F27D47CDA5AB9AAA883A66>.
- (1995b). *Educational attainments in developing countries: new estimates and projections disaggregated by gender*. Policy Research Working Paper Series 1489. The World Bank. URL: <http://elibrary.worldbank.org/content/workingpaper/10.1596/1813-9450-1489>.
- Alachkar, A. e W. J. Serow (1988). The socioeconomic determinants of mortality: An international comparison. *Genus* 44.3-4, 131–151.
- Alders, M. et al. (2007). Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population/Revue européenne de Démographie* 23 (1). 10.1007/s10680-006-9104-4, 33–69. ISSN: 0168-6577. URL: <http://dx.doi.org/10.1007/s10680-006-9104-4>.
- Alho, J. (2006). New forecast: Population decline postponed in Europe. *Statistical Journal of the United Nations ECE* 23.1-10.
- Alho, J. e N. Keilman (2010). On future household structure. *Journal of the Royal Statistical Society* 173, 117–143.
- Alho, J. M. et al. (2008). “Empirically based specification of forecast uncertainty”. Em: *Uncertain Demographics and Fiscal Sustainability*. Cambridge University Press. URL: <http://dx.doi.org/10.1017/CB09780511493393>.
- Alkema, L. et al. (2011). Probabilistic Projections of the Total Fertility Rate for All Countries. *Demography* 48 (3). 10.1007/s13524-011-0040-5, 815–839. ISSN: 0070-3370. URL: <http://dx.doi.org/10.1007/s13524-011-0040-5>.
- Anderson, T. W. e L. A. Goodman (1957). Statistical Inference about Markov Chains. English. *The Annals of Mathematical Statistics* 28.1, pp. 89–110. ISSN: 00034851. URL: <http://www.jstor.org/stable/2237025>.

- Angrist, J. (2006). *New evidence on the causal link between the quantity and quality of children*. Rel. téc. IZA Discussion Paper, No. 2075, April 2006. Bonn.
- BancoMundial (2012). *Education: Overview*. Acedido em: 12/04/2012. URL: <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,menuPK:282393~pagePK:162100~piPK:159310~theSitePK:282386,00.html>.
- Barreto, A. (2004). *Sociedade, demografia e políticas de população*. II Congresso Português de Demografia. Associação Portuguesa de Demografia. URL: http://www.apdemografia.pt/ficheiros_comunicacoes/1417468115.pdf.
- Barro, R. J. e J. Lee (1993). *International Comparisons of Educational Attainment*. NBER Working Papers 4349. National Bureau of Economic Research, Inc. URL: http://www.nber.org/papers/w4349.pdf?new_window=1.
- (2000). *International Data on Educational Attainment Updates and Implications*. NBER Working Papers 7911. National Bureau of Economic Research, Inc. URL: http://www.economics.harvard.edu/faculty/barro/files/p_jwha.pdf.
- Bernardelli, H. (1941). Population waves. *Journal of the Burma Research Society* 31, 1–18.
- Billari, F. C. e D. Philipov (2004). *Education and the Transition to Motherhood: a Comparative Analysis of Western Europe*. Vienna Institute of Demography Publications. URL: http://www.oew.ac.at/vid/download/edrp_3_04.pdf.
- Birdsall, N.M. (1988). Fertility and poverty in developing countries. *Journal of Policy Modeling* 10.1, 29–55. DOI: 10.1016/0161-8938(88)90034-8.
- Brass, W. (1971). “On the Scale of Mortality”. Em: *Biological Aspects of Demography*. Ed. por W. Brass. London: Taylor & Francis.
- Brewster, K. L. (2000). Fertility and women’s employment in industrialized nations. *Annual Review of Sociology* 26, 271–296. DOI: 10.1146/annurev.soc.26.1.271.
- California, University of e Max Planck Institute for Demographic Research (2009). *Human Mortality Database*. URL: <http://www.mortality.org>.
- Cannan, E. (1895). The probability of a cessation of the growth of population in England and Wales during the next century. *Economic Journal* 5.20, 505–15. DOI: 10.2307/2956626.
- Carrilho, M. J. (2005). Metodologias de cálculo das projecções demográficas: aplicação em Portugal. *Revista de Estudos Demográficos, INE, I.P.* 37.1, 5–24. URL: http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=378782&PUBLICACOESmodo=2.
- Carrilho, M. J. e L. Patrício (2010). A Situação Demográfica Recente em Portugal. *Revista de Estudos Demográficos, INE, I.P.* 48.5, 101–138. URL: http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_estudos&ESTUDOSest_boui=103155854&ESTUDOSmodo=2.
- Carvalho, M. L. (2010). Notas teóricas da disciplina de Demografia do Mestrado em Bioestatística da Faculdade de Ciências da Universidade de Lisboa (ano lectivo 2010/2011).

- Clark, D. e H. Royer (2010). *The Effect of Education on Adult Health and Mortality: Evidence from Britain*. NBER Working Paper No. 16013. World Bank, Washington, DC.
- Clark, S. J. e D. J. Sharrow (2011). *Contemporary Model Life Tables for Developed Countries: An Application of Model-based Clustering*. Working Paper 107. Center for Statistics e the Social Sciences. University of Washington, Seattle. URL: <http://www.csss.washington.edu/Papers/wp107.pdf>.
- Cleland, J. (2002). "Education and future fertility trends, with special reference to mid-transitional countries". Em: *Completing the Fertility Transition. Part 2. Background Papers*. New York: United Nations, Population Division, Department of Economic e Social Affairs, pp. 187–202.
- Coale, A. (1966). *Regional model life tables and stable populations*. Princeton: Princeton University Press.
- Cox, D.R. e H.D. Miller (1965). *The theory of stochastic processes*. Wiley publications in statistics. Wiley.
- Crowder, M. e D. Stephens (2011). On inference from Markov chain macro-data using transforms. *Journal of Statistical Planning and Inference* 141.9, 3201–3216. ISSN: 0378-3758. DOI: 10.1016/j.jspi.2011.04.007. URL: <http://www.sciencedirect.com/science/article/pii/S0378375811001509>.
- Davis, B. A. et al. (2002). Estimating and Interpolating a Markov Chain from Aggregate Data. English. *Biometrika* 89.1, pp. 95–110. ISSN: 00063444. URL: <http://www.jstor.org/stable/4140561>.
- Davis, B.A. et al. (2001). Estimating cohort health expectancies from cross-sectional surveys of disability. *Statistics in Medicine* 20, 1097–1111. DOI: 10.1002/sim.724.
- Decreto-Lei, n.º 301/93, de 31 de Agosto. *Diário da República, 1.ª série-A — N.º 204*. Ministério da Educação. Lisboa. URL: <http://dre.pt/pdf1sdip/1993/08/204A00/45934599.pdf>.
- Despacho, Normativo n.º 1/2005, de 5 de Janeiro. *Diário da República, 1.ª série-B — N.º 3*. Ministério da Educação. Lisboa. URL: <http://dre.pt/pdf1sdip/2005/01/003B00/00710076.pdf>.
- Despacho, Normativo n.º 29/2010, de 14 de Dezembro. *Diário da República, 2.ª série — N.º 240*. Ministério da Educação, Gabinete do Secretário de Estado Adjunto e da Educação. Lisboa. URL: <http://dre.pt/pdf2sdip/2010/12/240000000/6035060351.pdf>.
- Dorn, H. F. (1950). Pitfalls in Population Forecasts and Projections. English. *Journal of the American Statistical Association* 45.251, pp. 311–334. ISSN: 01621459. URL: <http://www.jstor.org/stable/2280290>.
- Dustmann, C. e A. Glitz (2011). *Migration and Education*. Norface Discussion Paper Series 2011011. Norface Research Programme on Migration, Department of Economics, University College London. URL: http://www.norface-migration.org/publ_uploads/NDP_11_11.pdf.
- Eurostat (2010). *Regional population projections EUROPOP2008: Most EU regions face older population profile in 2030*. Acedido em: 05/04/2012. URL: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-10-001/EN/KS-SF-10-001-EN.PDF.
- (2012). *Statistics Database*. Acedido em: 29/03/2012. URL: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

- Fernandes, A. A. (2007). Determinantes da mortalidade e da longevidade: Portugal numa perspectiva europeia (UE15, 1991-2001). *Análise Social* XLII.183, 419–443. URL: <http://www.scielo.gpeari.mctes.pt/pdf/aso/n183/n183a03.pdf>.
- GEPE (2012). *Sistema Educativo - Apresentação do Sistema Educativo Português*. Gabinete de Estatística e Planeamento da Educação, Ministério da Educação. Acedido em: 06/03/2012. URL: <http://www.gepe.min-edu.pt/np4/9.html>.
- Gill, R. e N. Keilman (1990). On the estimation of multidimensional demographic models with population registration data. *Mathematical Population Studies* 2.2, 119–143. DOI: 10.1080/08898489009525298. URL: <http://oai.cwi.nl/oai/asset/5926/5926A.pdf>.
- Goujon, A. (2008). *Report on projections by level of education (Future human capital: Estimates and projections of education transition probabilities)*. Deliverable D3 in Work Package 1. Bridging the micro-macro gap in population forecasting - MicMac Project. Vienna Institute of Demography, Austrian Academy of Sciences. URL: <http://www.nidi.nl/Content/NIDI/output/micmac/micmac-d3.pdf>.
- Goujon, A. e A. Wils (1996). *The Importance of Education in Future Population. Global Trends and Case Studies on Cape Verde, Sudan, and Tunisia*. Working Paper WP-96-138. International Institute for Applied Systems Analysis. Laxenburg, Austria. URL: <http://www.iiasa.ac.at/Admin/PUB/Documents/WP-96-138.pdf>.
- Goujon, A. et al. (2007). *Population and Human Capital Growth in Egypt: Projections for Governance rates to 2051*. Interim Report IR-07-010. International Institute for Applied Systems Analysis. Laxenburg, Austria. URL: <http://www.iiasa.ac.at/Research/POP/pdes/egypt/docs/IR-07-010.pdf>.
- Gouno, E. et al. (2011). Estimation from aggregate data. *Comput. Stat. Data Anal.* 55.1, 615–626. ISSN: 0167-9473. DOI: 10.1016/j.csda.2010.06.003. URL: <http://dx.doi.org/10.1016/j.csda.2010.06.003>.
- Gustavsson, S. (2006). *Education and Postponement of Maternity. Economic Analyses for Industrialized Countries*. Dordrecht, The Netherlands: Kluwer Academic Publishers e Springer. DOI: 10.1007/1-4020-4716-9.
- Haddad, W. D. et al. (1990). *Education and Development: Evidence for New Priorities*. World Bank Discussion Papers No. 95. World Bank, Washington, DC. URL: <http://www.eric.ed.gov/PDFS/ED326471.pdf>.
- Heathcote, C. R. e B. D. Puza (2005). Estimating processes of counts from cross-sectional aggregate data, with an application to multistate life tables and health expectancies. *Australian & New Zealand Journal of Statistics* 47.4, 425–439. ISSN: 1467-842X. DOI: 10.1111/j.1467-842X.2005.00406.x. URL: <http://dx.doi.org/10.1111/j.1467-842X.2005.00406.x>.
- Henriques, F. C. e T. F. Rodrigues (2010). “Essay on Ageing and Health projections in Portugal”. Em: *Item 3 – Challenges and use of population projections*. Joint Eurostat/UNECE Work Session on

- Demographic Projections (28-30 April 2010, Lisbon, Portugal). Statistical Office of the European Union (EUROSTAT).
- Hougaard, P. (1999). Multi-state Models: A Review. *Lifetime Data Analysis* 5 (3), 239–264. ISSN: 1380-7870. URL: <http://dx.doi.org/10.1023/A:1009672031531>.
- Imhoff, E. e N. Keilman (1991). *Lipro 2.0: an application of a dynamic demographic projection model to household structure in the Netherlands*. Lipro 2.0. Swets & Zeitlinger. ISBN: 9789026512414. URL: <http://nidi.nl/Content/NIDI/output/lipro/nidicbgs-publ-23.pdf>.
- INE (2008). *Projeções de população residente em Portugal - 2008-2060 - Metodologia*. Estudo. URL: http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_estudos&ESTUDOSest_boui=65946997&ESTUDOSmodo=2.
- (2010). *Documento Metodológico - Tábuas Completas de Mortalidade*. Versão 1.0, Código 84. Departamento de Estatísticas Demográficas e Sociais, Serviço de Estatísticas Demográficas. URL: http://metaweb.ine.pt/sim/OPERACOES/DOCMET_PDF/DOCMET_PDF_84_1_0.pdf.
- (2012). *Metainformação*. Acedido em: 30/03/2012. URL: <http://metaweb.ine.pt/sim/conceitos/conceitos.aspx>.
- INE, IP, ed. (2009). *Projeções de população residente em Portugal 2008-2060*. Lisboa, Portugal: Instituto Nacional de Estatística, IP.
- Jain, A. K. (1981). The effect of female education on fertility: A simple explanation. *Demography* 18.4, 577–595. DOI: 10.2307/2060948.
- Jejeebhoy, S.J. (1995). *Women's Education, Autonomy and Reproductive Behaviour: Experience from Developing Countries*. Oxford:Clarendon Press.
- Jones, M. T. (2005). *Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk*. IMF Working Papers 05/219. International Monetary Fund. URL: <http://ideas.repec.org/p/imf/imfwpa/05-219.html>.
- Kalbfleisch, J. D. e J. F. Lawless (1984). Least-squares estimation of transition probabilities from aggregate data. *Canadian Journal of Statistics* 12.3, 169–182. ISSN: 1708-945X. DOI: 10.2307/3314745. URL: <http://dx.doi.org/10.2307/3314745>.
- Kalbfleisch, J. D. et al. (1983). Estimation in Markov Models from Aggregate Data. English. *Biometrics* 39.4, pp. 907–919. ISSN: 0006341X. URL: <http://www.jstor.org/stable/2531326>.
- KC, S. et al. (2010). Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research* 22.15, 383–472.
- Ledermann, S. (1969). *Nouvelles tables-types de mortalité*. Travaux et documents - Institut national d'études démographiques. Presses Universitaires de France.
- Lei, n.º 85/2009, de 27 de Agosto. *Diário da República*, 1.ª série — N.º 166. Assembleia da República. Lisboa. URL: <http://dre.pt/pdf1sdip/2009/08/16600/0563505636.pdf>.
- Leslie, P.H. (1945). On the use of matrices in certain population mathematics. *Biometrika* 33, 183–212.
- Lewis, E.G. (1942). On the generation and growth of a population. *Sankhya* 6, 93–96.

- Liu, X. (2010). “Ordinal Regression Analysis: Fitting the Continuation Ratio Model to Educational Data Using Stata”. Em: *Paper 35*. NERA Conference Proceedings 2010. URL: http://digitalcommons.uconn.edu/nera_2010/35.
- Lleras-Muney, A. (2005). The Relationship between Education and Adult Mortality in the United States. *The Review of Economic Studies* 72.1, 189–221.
- Lutz, W. (1999). “Demographic dimensions in forecasting: Adding education to age and sex”. Em: *Frontiers of Population Forecasting. Supplement to Volume 24, 1998, Population and Development Review*. Ed. por W. Lutz. New York: The Population Council, pp. 42–58.
- Lutz, W. e A. Goujon (2001). The World’s Changing Human Capital Stock: Multi-State Population Projections by Educational Attainment. English. *Population and Development Review* 27.2, pp. 323–339. ISSN: 00987921. URL: <http://www.jstor.org/stable/2695213>.
- Lutz, W. e S. KC (2010). Dimensions of global population projections: what do we know about future population trends and structures? *Philosophical Transactions of the Royal Society B* 365, 2779–2791.
- Lutz, W. et al. (1997). Doubling of world population unlikely. *Nature* 387, 803–805. DOI: 10.1038/42935.
- (2001). The end of world population growth. *Nature* 412, 543–545. DOI: 10.1038/35087589.
- Lutz, W. et al. (2005). *Forecasting Human Capital: Using Demographic Multi-State Methods by Age, Sex, and education to Show the Long-Term Effects of Investments in Education*. Working Paper WP-07-03. Washington DC: Education Policy e Data Center, Academy for Educational Development. URL: <http://paa2005.princeton.edu/download.aspx?submissionId=50235>.
- Lutz, W. et al. (2008). The coming acceleration of global population ageing. *Nature* 451, 716–719. DOI: 10.1038/nature06516.
- Magalhães, M. G. (2003). Migrações inter NUTS II e projecções regionais de População Residente. *Revista de Estudos Demográficos, INE, I.P.* 34.5, 61–71. URL: http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=378613&PUBLICACOESmodo=2.
- Mahlmann, T. (2006). Estimation of rating class transition probabilities with incomplete data. *Journal of Banking & Finance* 30.11, 3235–3256. URL: http://www.ku.de/fileadmin/160115/Publikationen/Abstract/Estimation_of_rating_class_transition.pdf.
- McQuillan, K. (2004). When does religion influence fertility? *Population and Development Review* 30.1, 25–56. DOI: 10.1111/j.1728-4457.2004.00002.x.
- ME. *Programa Educação 2015*. Ministério da Educação. URL: <http://www.drelvt.min-edu.pt/alunos/programa-educacao-2015.pdf>.
- (2012). *Ministério da Educação - Legislação sobre Educação*. Acedido em: 20/02/2012. URL: <http://legislacao.min-edu.pt/np4/133>.
- Müller, D. (2007). *Processos estocásticos e aplicações*. Ed. por Edições Almedina. Coleção Económicas II Série, N.º3. Coimbra, Portugal.

- Moffitt, R. (1993). Identification and estimation of dynamic models with a time series of repeated cross-sections. *Journal of Econometrics* 59.1-2, 99–123. URL: http://www.uio.no/studier/emner/sv/oekonomi/ECON5103/v10/undervisningsmateriale/PDApl_11.pdf.
- MS, ed. (2009). *SNS 30 Anos - Indicadores*. Ministério da Saúde. URL: <http://www.sns30anos.min-saude.pt/SiteCollectionDocuments/SNSIndicadores.pdf>.
- Murray, C. J. L. et al. (2000). *WHO System of Model Life Tables*. GPE discussion paper 8. URL: <http://www.who.int/healthinfo/paper08.pdf>.
- Nehru, V. et al. (1995). A new database on human capital stock in developing and industrial countries: Sources, methodology, and results. *Journal of Development Economics* 46.2, 379–401. URL: http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/1993/04/01/000009265_3961004114011/Rendered/PDF/multi_page.pdf.
- Notestein, F. W. (1945). “Population - The Long View”. Em: *Food for the World*. Ed. por T. W. Schultz. Chicago, Illinois: University of Chicago Press, pp. 36–57.
- OCDE (2001). *Organisation for Economic Co-operation and Development. Glossary of Statistical Terms - Population Projections*. URL: <http://stats.oecd.org/glossary/detail.asp?ID=2088>.
- (2011). *Health at a Glance 2011: OECD Indicators*. OECD Publishing. URL: http://dx.doi.org/10.1787/health_glance-2011-en.
- Oliveira, I. T. (2008). Fecundidade das populações e das gerações em Portugal, 1960-2005. *Análise Social* XLIII.1, 29–53. URL: <http://www.scielo.gpeari.mctes.pt/pdf/aso/n186/n186a03.pdf>.
- (2009). Fertility Differentials and Educational Attainment in Portugal: A Non-Linear Relationship. *Canadian Studies in Population* 36.3-4, 347–362. URL: <http://www.canpopsoc.org/journal/2009/CSPv36n3-4p347.pdf>.
- Oliveira, I. T. e M. F. Mendes (2010). A diferença de esperança de vida entre homens e mulheres: Portugal de 1940 a 2007. *Análise Social* XLV.194, 115–138. URL: <http://analisesocial.ics.ul.pt/documentos/1268307381P1wGC4sr3Qg28GG4.pdf>.
- O’Neill, B. C. et al. (2001). A Guide to Global Population Projections. *Demographic Research* 4.8, 203–288.
- ONU (2011a). *United Nations Department of Economic and Social Affairs Population Division. Assumptions underlying the 2010 revision. Department of Economic and Social Affairs, Population Division*. URL: http://esa.un.org/wpp/Documentation/pdf/WPP2010_ASSUMPTIONS_AND_VARIANTS.pdf.
- (2011b). *United Nations Department of Economic and Social Affairs Population Division. World Population Prospects: The 2010 Revision, Highlights and Advance Tables*. URL: <http://esa.un.org/wpp/Documentation/publications.htm>.
- Pedrosa, A. P. (2011). “Modelação e projecção estocástica da população portuguesa para 2050”. Tese de mestrado. Departamento de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.

- Pelzer, B. et al. (2001). Estimating Transition Probabilities from a Time Series of Independent Cross Sections. *Statistica Neerlandica* 55.2, 249–262. ISSN: 1467-9574. DOI: 10.1111/1467-9574.00168. URL: <http://dx.doi.org/10.1111/1467-9574.00168>.
- Psacharopoulos, G. e A.M. Arriagada (1992). *The Educational Composition of the Labor Force: An International Update*. PHREE background paper series. World Bank.
- Psacharopoulos, G. e H. A. Patrinos (2004). Returns to Investment in Education: A Further Update. *Education Economics* 12.2, 111–134.
- Putter, H. et al. (2007). Tutorial in biostatistics: competing risks and multi-state models. eng. *Stat. Med.* 26.11. CI: Copyright 2006; JID: 8215016; ppublish, 2389–2430.
- Resolução, do Conselho de Ministros n.º 44/2010, de 14 de Junho. *Diário da República*, 1.ª série — N.º 113. Presidência do Conselho de Ministros. Lisboa. URL: <http://dre.pt/pdf1sdip/2010/06/11300/0199701998.pdf>.
- Rogers, A. (1975). *Introduction to multiregional mathematical demography*. A Wiley-Interscience publication. Wiley. ISBN: 9780471730354.
- Ryder, N. B. (1965). The Cohort as a Concept in the Study of Social Change. *American Sociological Review* 30.6, 843–861. URL: <http://www.jstor.org/discover/10.2307/2090964?uid=3738880&uid=2129&uid=2134&uid=2&uid=70&uid=4&sid=47699104670467>.
- Scherbov, S. e M. Mamolo (2006). *Probabilistic Population Projections for the EU-25*. European Demographic Research Papers 1. Vienna Institute of Demography of the Austrian Academy of Sciences. URL: http://www.oeaw.ac.at/vid/download/edrp_1_06.pdf.
- Scherbov, S. et al. (2008). *Probabilistic Population Projections for the 27 EU Member States Based on Eurostat Assumptions*. European Demographic Research Papers 1. Vienna Institute of Demography of the Austrian Academy of Sciences. URL: http://www.oeaw.ac.at/vid/download/edrp_2_08.pdf.
- Schoen, R. (2006a). “Dynamic Birth-Death (Multi-Age) Models”. Em: *Dynamic Population Models*. Vol. 17. The Springer Series on Demographic Methods and Population Analysis. Springer Netherlands, pp. 123–160. ISBN: 978-1-4020-5230-9.
- (2006b). “Dynamic Models With Multiple Ages and States”. Em: *Dynamic Population Models*. Vol. 17. The Springer Series on Demographic Methods and Population Analysis. Springer Netherlands, pp. 203–230. ISBN: 978-1-4020-5230-9.
- (2006c). “Dynamic Multistate Models”. Em: *Dynamic Population Models*. Vol. 17. The Springer Series on Demographic Methods and Population Analysis. Springer Netherlands, pp. 161–202. ISBN: 978-1-4020-5230-9.
- (2006d). “Population Models With Constant Rates”. Em: *Dynamic Population Models*. Vol. 17. The Springer Series on Demographic Methods and Population Analysis. Springer Netherlands, pp. 1–26. ISBN: 978-1-4020-5230-9.
- Sharpe, F. R. e A. J. Lotka (1911). A problem in age-distribution. *Philosophical Magazine* 21 (124), 435–438.

- Sharrow, D. J. e H. GUI by Sevcikova (2011). *LifeTables: A package to implement HMD model life table system*. R package version 0.1. URL: <http://CRAN.R-project.org/package=LifeTables>.
- Shryock, H.S. et al. (1973). *The Methods and Materials of Demography*. The Methods and Materials of Demography vol. 1. U.S. Department of Commerce, Social & Economic Statistics Administration, Bureau of the Census.
- Siegel, J. S. (2012). “Models of Aging, Health, and Mortality, and Mortality/Health Projections”. Em: *The Demography and Epidemiology of Human Health and Aging*. Springer Netherlands, pp. 731–773. ISBN: 978-94-007-1315-4.
- Skirbekk, V. (2008). Fertility trends by social status. *Demographic Research* 18.5, 145–180. DOI: 10.4054/DemRes.2008.18.5.
- Skirbekk, V. et al. (2007). *Report on methods for demographic projections at multiple levels of aggregation*. D1.2.1. PLUREL Project. IIASA, Laxenburg. URL: http://www.iiasa.ac.at/Research/POP/pub/Skirbekk/Plurel-D121_Skirbekk_etal.pdf.
- Sleebos, J. (2003). *Low Fertility Rates in OECD Countries: Facts and Policy Responses*. OECD Labour Market and Social Policy Occasional Papers 15. OECD Publishing. URL: <http://ideas.repec.org/p/oec/elsaaa/15-en.html>.
- Testa, M.R. (2006). The influence of childbearing regional contexts on ideal family size in Europe. *Population* 61.1-2, 109–138.
- UE. *EF 2020*. Conclusões do Conselho de 12 de Maio de 2009 sobre um quadro estratégico para a cooperação europeia no domínio da educação e da formação («EF 2020»), Jornal Oficial da União Europeia C 119 de 28.5.2009. URL: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:119:0002:0010:PT:PDF>.
- UN (1982). *Model Life Tables for Developing Countries*. United Nations publication Sales No. E.81.XIII.7. URL: http://www.un.org/esa/population/publications/Model_Life_Tables/Model_Life_Tables.htm.
- (1995). *Women’s Education & Fertility Behaviour: Recent Evidence from the Demographic and Health Surveys*. New York: United Nations.
- (2011). *Trends in International Migrant Stock: Migrants by Age and Sex*. United Nations, Department of Economic and Social Affairs, Population Division. URL: <http://esa.un.org/MigAge/index.asp?panel=1>.
- UNESCO (2006). *ISCED 1997 - International Standard Classification of Education*. Re-edition. URL: <http://www.uis.unesco.org/Library/Documents/isced97-en.pdf>.
- Van Peer, C. (2002). “Desired and realized fertility in selected FFS-countries”. Em: *Dynamics of Fertility and Partnership in Europe: Insights and Lessons from Comparative Research, Vol. 1*. Ed. por M. Macura. New York e Geneva: United Nations, pp. 117–142.
- WHO (2012). *Definition of an older or elderly person*. URL: <http://www.who.int/healthinfo/survey/ageingdefnolder/en/index.html>.

- Willekens, F. (2006). *Description of the multistate projection model (Multistate model for biographic analysis and projection)*. Deliverable D1 in Work Package 1. Bridging the micro-macro gap in population forecasting - MicMac Project. Netherlands Interdisciplinary Demographic Institute. The Hague, The Netherlands. URL: <http://www.nidi.nl/Content/NIDI/output/micmac/micmac-d1.pdf>.
- Willekens, F. J. (1990). "Demographic forecasting. State-of-the-art and research needs". Em: *Emerging issues in demographic research*. Ed. por C.A. Hazeu e G.A.B. Frinking. Elsevier Science Publishers B.V. Cap. 2, pp. 9-66. URL: <http://depot.knaw.nl/2448/1/18930.pdf>.
- Wils, A. (2007). *Window on the Future: 2025 - Projections of Education Attainment and Its Impact*. Washington, DC: Education Policy e Data Center. URL: <http://epdc.org/policyanalysis/static/WindowOnTheFuture2025.pdf>.
- Wils, A. et al. (2005). *Educating the World's Children: Patterns of Growth and Inequality*. Washington, DC: Education Policy e Data Center. URL: http://pdf.usaid.gov/pdf_docs/PNADL901.pdf.
- Wilson, T. e P. Rees (2005). Recent developments in population projection methodology: A review. *Population, Space and Place* 11, 337-360.

APÊNDICE A

Tabelas auxiliares

TABELA A.0.1. *População residente (N.º) por sexo e grupo etário, 31/12/2000*

Fonte: INE, *Estimativas Anuais da População Residente*

	Grupo etário (anos)																
Sexo	Total	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
H	4.950.696	273.233	272.503	293.440	347.923	396.131	405.224	375.615	375.012	353.995	330.086	306.426	266.241	253.650	241.812	194.673	264.732
M	5.305.962	261.053	259.891	280.555	334.087	387.174	401.656	378.529	388.319	367.535	349.457	329.927	299.724	291.985	291.188	254.947	429.935
HM	10.256.658	534.286	532.394	573.995	682.010	783.305	806.880	754.144	763.331	721.530	679.543	636.353	565.965	545.635	533.000	449.620	694.667

TABELA A.0.2. *Distribuição dos residentes de cada sexo e grupo etário por nível de escolaridade, Censos 2001*

Fonte: INE, *XIV Recenseamento Geral da População*

Nível de escolaridade	Sexo	Grupo etário (anos)															
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
SE / 2.ºCiclo EB	H	1,000	1,000	0,984	0,391	0,322	0,418	0,519	0,581	0,621	0,684	0,716	0,752	0,805	0,840	0,860	0,878
	M	1,000	1,000	0,984	0,280	0,234	0,350	0,473	0,555	0,616	0,702	0,752	0,814	0,870	0,893	0,910	0,929
Ensino Básico	H	0,000	0,000	0,016	0,525	0,313	0,222	0,199	0,177	0,149	0,109	0,105	0,096	0,076	0,063	0,052	0,043
	M	0,000	0,000	0,016	0,585	0,249	0,167	0,165	0,153	0,134	0,097	0,089	0,071	0,051	0,042	0,036	0,030
Ensino Secundário	H	0,000	0,000	0,000	0,085	0,329	0,235	0,173	0,150	0,141	0,114	0,102	0,089	0,070	0,060	0,053	0,045
	M	0,000	0,000	0,000	0,135	0,435	0,256	0,188	0,155	0,136	0,100	0,082	0,061	0,045	0,042	0,037	0,029
Ensino Superior	H	0,000	0,000	0,000	0,000	0,036	0,125	0,110	0,092	0,089	0,093	0,077	0,063	0,048	0,037	0,035	0,034
	M	0,000	0,000	0,000	0,000	0,082	0,227	0,175	0,137	0,113	0,101	0,076	0,054	0,035	0,023	0,017	0,012

TABELA A.0.3. *População de partida por sexo, grupo etário e nível de escolaridade (1 de janeiro de 2001)*

Fonte: *Cálculos do autor*

Nível de escolaridade	Sexo	Grupo etário (anos)															
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
SE / 2.ºCiclo EB	H	273.233	272.503	288.882	135.885	127.423	169.220	194.837	217.875	219.754	225.638	219.428	200.293	204.210	203.168	167.331	232.474
	M	261.053	259.891	276.042	93.541	90.520	140.507	179.071	215.587	226.450	245.395	248.197	243.895	253.935	260.035	232.088	399.343
Ensino Básico	H	0	0	4.558	182.513	123.900	89.870	74.672	66.389	52.816	36.055	32.123	25.506	19.363	15.160	10.213	11.424
	M	0	0	4.513	195.429	96.514	67.083	62.336	59.250	49.251	34.019	29.388	21.297	14.791	12.241	9.078	12.751
Ensino Secundário	H	0	0	0	29.524	130.357	95.316	64.964	56.311	50.002	37.700	31.229	23.777	17.840	14.449	10.288	11.902
	M	0	0	0	45.117	168.341	102.837	71.050	60.161	50.123	34.789	27.105	18.238	13.112	12.146	9.446	12.675
Ensino Superior	H	0	0	0	0	14.451	50.819	41.142	34.437	31.423	30.694	23.645	16.665	12.237	9.035	6.841	8.932
	M	0	0	0	0	31.799	91.230	66.072	53.321	41.711	35.254	25.236	16.294	10.147	6.766	4.335	5.166

TABELA A.0.4. *Taxa de mortalidade infantil (‰) por nível de escolaridade da mãe, 2000-2007*

Fonte: *Cálculos do autor*

Ano	Nível de escolaridade			
	SE / 2.ºCiclo EB	Ensino Básico	Ensino Secundário	Ensino Superior
2000	6,31	6,32	4,38	2,30
2001	5,94	4,98	3,99	2,62
2002	6,50	5,49	3,95	1,76
2003	5,25	5,56	3,09	1,94
2004	4,78	4,25	3,10	1,84
2005	4,26	4,37	2,64	1,35
2006	4,04	4,10	3,18	1,57
2007	4,97	3,51	2,98	2,26

TABELA A.0.5. *Quociente de mortalidade entre as idades exatas x e $(x + n)$ - ${}_nq_x$ - por sexo e grupo etário, 2000-2002 a 2005-2007*

Fonte: INE, *Tábuas Completas de Mortalidade*

Sexo	Grupo etário (anos)	Período					
		2000-2002	2001-2003	2002-2004	2003-2005	2004-2006	2005-2007
H	0	0,0058	0,0052	0,0046	0,0043	0,0039	0,0036
	1-4	0,0018	0,0016	0,0013	0,0013	0,0012	0,0010
	5-9	0,0011	0,0011	0,0011	0,0011	0,0010	0,0008
	10-14	0,0014	0,0013	0,0012	0,0010	0,0010	0,0009
	15-19	0,0042	0,0039	0,0034	0,0032	0,0033	0,0029
	20-24	0,0068	0,0060	0,0053	0,0050	0,0045	0,0042
	25-29	0,0088	0,0076	0,0066	0,0061	0,0054	0,0048
	30-34	0,0107	0,0099	0,0091	0,0085	0,0078	0,0070
	35-39	0,0134	0,0131	0,0120	0,0115	0,0113	0,0106
	40-44	0,0178	0,0181	0,0169	0,0163	0,0157	0,0154
	45-49	0,0244	0,0243	0,0238	0,0230	0,0223	0,0223
	50-54	0,0309	0,0318	0,0312	0,0313	0,0309	0,0308
	55-59	0,0456	0,0453	0,0441	0,0433	0,0431	0,0426
	60-64	0,0691	0,0678	0,0657	0,0649	0,0621	0,0598
	65-69	0,1085	0,1069	0,1032	0,0999	0,0972	0,0945
	70-74	0,1769	0,1707	0,1643	0,1622	0,1551	0,1498
	75-79	0,2769	0,2769	0,2709	0,2685	0,2579	0,2491
	80-84	0,4341	0,4319	0,4208	0,4191	0,4094	0,4056
	85-89	0,7063	0,6976	0,6895	0,6688	0,6496	0,6808
	90-94	0,8956	0,8907	0,8874	0,8747	0,8618	0,8830
	95-99	0,9791	0,9785	0,9772	0,9736	0,9701	0,9760
	100 +	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
M	0	0,0045	0,0046	0,0044	0,0036	0,0034	0,0030
	1-4	0,0014	0,0012	0,0010	0,0010	0,0009	0,0008
	5-9	0,0010	0,0009	0,0007	0,0008	0,0008	0,0006
	10-14	0,0010	0,0010	0,0009	0,0009	0,0006	0,0007
	15-19	0,0015	0,0013	0,0015	0,0015	0,0013	0,0011
	20-24	0,0017	0,0018	0,0018	0,0017	0,0015	0,0014
	25-29	0,0023	0,0024	0,0022	0,0021	0,0018	0,0016
	30-34	0,0033	0,0032	0,0031	0,0030	0,0027	0,0024
	35-39	0,0047	0,0044	0,0043	0,0043	0,0040	0,0039
	40-44	0,0075	0,0069	0,0064	0,0063	0,0059	0,0058
	45-49	0,0105	0,0104	0,0098	0,0095	0,0091	0,0089
	50-54	0,0140	0,0142	0,0140	0,0136	0,0131	0,0123
	55-59	0,0205	0,0202	0,0199	0,0198	0,0185	0,0172
	60-64	0,0323	0,0316	0,0303	0,0288	0,0277	0,0273
	65-69	0,0515	0,0501	0,0486	0,0486	0,0468	0,0439
	70-74	0,0972	0,0940	0,0897	0,0874	0,0829	0,0785
	75-79	0,1782	0,1753	0,1711	0,1671	0,1567	0,1488
	80-84	0,3238	0,3216	0,3117	0,3078	0,2959	0,2900
	85-89	0,5889	0,5860	0,5660	0,5515	0,5348	0,5604
	90-94	0,8297	0,8279	0,8150	0,8045	0,7910	0,8132
	95-99	0,9609	0,9605	0,9565	0,9530	0,9487	0,9558
	100 +	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

TABELA A.0.6. População feminina com 15-49 anos (milhares de indivíduos), por grupo etário e nível de escolaridade, 1998-2010

Fonte: INE, Estatísticas do Emprego

Nível de escolaridade	Ano	Grupo etário (anos)						
		15-19	20-24	25-29	30-34	35-39	40-44	45-49
SE / 2.ºCiclo EB	1998	124,5	140,3	187,5	225,4	236,1	258,5	264,3
	1999	119,6	124,2	185,3	218,2	242,5	245,9	265,5
	2000	105,1	101,4	183,5	210,1	243,7	247,1	264,0
	2001	98,2	86,5	169,8	204,8	239,2	249,0	264,2
	2002	88,5	76,8	153,3	204,6	235,4	253,5	265,2
	2003	74,6	74,0	138,9	200,8	227,4	246,4	253,9
	2004	69,3	55,7	131,4	179,8	220,1	232,4	241,1
	2005	71,3	51,6	107,9	172,6	213,6	230,5	235,9
	2006	70,6	43,9	99,9	160,4	194,7	227,2	227,7
	2007	69,8	42,2	90,6	154,6	190,6	224,2	231,5
	2008	58,5	35,8	61,5	134,2	173,7	217,0	239,2
	2009	47,8	34,1	52,6	113,8	157,0	202,8	223,0
	2010	48,3	26,7	45,6	95,1	142,4	195,0	213,8
Ensino Básico	1998	193,2	79,8	56,0	53,9	49,8	33,3	28,8
	1999	194,1	83,1	53,8	53,2	48,7	39,5	29,3
	2000	188,5	86,0	50,6	55,9	51,2	44,3	35,4
	2001	182,8	93,7	57,4	53,3	55,9	49,8	33,5
	2002	184,4	100,5	60,4	55,1	55,9	52,7	33,2
	2003	186,0	90,5	68,1	52,5	52,1	54,1	40,4
	2004	174,4	91,5	71,0	60,3	54,0	62,3	44,0
	2005	169,6	99,2	76,2	64,2	54,9	61,3	49,2
	2006	166,7	98,0	77,8	66,5	56,6	64,4	58,1
	2007	162,2	83,8	76,9	74,5	67,8	71,1	59,5
	2008	174,4	82,5	91,3	88,5	75,8	75,3	64,4
	2009	180,6	79,2	93,3	105,5	76,4	78,1	72,3
	2010	172,8	75,6	82,0	90,7	80,4	75,2	82,1
Ensino Secundário	1998	39,5	153,9	79,7	50,6	45,9	28,8	18,9
	1999	35,6	159,5	87,8	55,3	48,5	35,1	18,3
	2000	41,6	175,7	92,1	60,9	44,4	35,8	19,5
	2001	43,5	174,9	92,9	65,0	46,2	35,2	23,6
	2002	41,0	166,1	100,7	62,1	49,3	39,4	24,5
	2003	41,9	171,4	107,6	67,6	49,7	44,1	25,6
	2004	47,5	173,7	108,0	77,9	52,6	47,2	31,0
	2005	44,1	166,5	119,1	83,5	59,0	53,0	40,5
	2006	40,8	168,5	113,7	94,4	65,5	53,8	43,8
	2007	45,0	161,7	101,5	92,7	63,3	46,1	45,6
	2008	46,1	152,7	98,5	92,8	69,8	45,6	44,2
	2009	49,3	142,8	96,3	100,1	82,7	56,1	46,3
	2010	52,5	143,6	107,9	108,6	91,5	59,1	49,8
Ensino Superior	1998	0,0	30,4	63,6	48,7	42,0	32,9	28,8
	1999	0,0	31,3	65,2	50,9	40,5	37,8	29,3
	2000	0,0	27,6	72,4	50,8	46,6	36,8	28,0
	2001	0,0	28,2	85,4	57,8	47,6	38,0	29,1
	2002	0,0	32,8	96,7	65,6	50,5	33,8	31,0
	2003	0,0	33,6	102,3	73,0	63,6	42,2	37,4
	2004	0,0	39,0	111,4	82,0	67,2	51,3	43,8
	2005	0,0	38,7	110,1	88,8	64,4	52,9	45,9
	2006	0,0	34,1	108,1	99,1	69,7	61,1	43,6
	2007	0,0	35,8	121,4	105,6	71,2	59,9	45,2
	2008	0,0	42,3	126,4	113,1	79,1	60,0	44,7
	2009	0,0	41,0	125,0	105,6	90,7	60,6	48,7
	2010	0,0	40,5	122,6	122,7	97,3	63,6	51,0

TABELA A.0.7. *Nados-vivos (N.^o) por grupo etário da mãe e nível de escolaridade da mãe, 1998-2010*Fonte: INE, *Nados-Vivos*

Nível de escolaridade	Ano	Grupo etário (anos)						
		15-19	20-24	25-29	30-34	35-39	40-44	45-49
SE / 2.ºCiclo EB	1998	4506	14193	17388	12833	5111	1150	71
	1999	4344	13365	17136	12655	5540	1243	83
	2000	4256	12446	16996	12597	5774	1293	91
	2001	3759	10456	14866	11696	5609	1247	81
	2002	3538	8989	13669	11170	5557	1218	94
	2003	3075	7232	11578	10177	5237	1248	70
	2004	2929	6242	10128	9318	4987	1180	95
	2005	2814	5620	8803	8881	4688	1189	80
	2006	2468	4962	7589	8073	4498	1219	64
	2007	2459	4410	6375	7146	4174	1071	50
	2008	2211	4073	5499	6426	4009	1029	67
	2009	1899	3588	4423	5354	3372	964	64
	2010	1757	3347	3937	4655	3217	936	56
Ensino Básico	1998	2001	5261	6405	4260	1619	231	11
	1999	2093	5427	6536	4177	1697	251	8
	2000	2218	5913	6644	4370	1834	329	20
	2001	2193	5653	6364	4328	1793	323	21
	2002	2264	5752	6637	4405	1822	395	15
	2003	2247	5983	6723	4596	1960	387	22
	2004	2204	5866	6807	4768	1990	449	25
	2005	1984	5919	6983	4770	2073	417	17
	2006	1748	5521	6515	4743	2019	406	25
	2007	1718	5248	6663	5061	2195	418	22
	2008	1679	5425	7132	5416	2437	513	24
	2009	1769	5085	6795	5318	2513	517	24
	2010	1641	4886	6663	5488	2644	585	30
Ensino Secundário	1998	793	4747	8301	6112	1991	231	9
	1999	819	5037	8955	6224	2292	337	11
	2000	899	5508	9474	6771	2530	344	19
	2001	831	4878	9130	6454	2375	380	20
	2002	836	5089	10065	6995	2624	438	19
	2003	745	5083	10408	7683	2758	460	17
	2004	607	4789	10121	7787	2870	522	32
	2005	645	4700	10257	8275	3187	548	25
	2006	594	4432	9779	8605	3326	586	19
	2007	558	4150	9097	8581	3459	599	18
	2008	531	4268	9290	9278	3767	647	28
	2009	573	4110	8918	9172	3988	715	34
	2010	490	4290	9395	9786	4403	789	26
Ensino Superior	1998	0	598	5804	6806	2541	292	7
	1999	0	708	6371	7267	2950	343	15
	2000	0	743	7013	8198	3132	426	27
	2001	0	739	7210	8374	3380	491	23
	2002	0	714	7903	9769	3745	527	24
	2003	0	655	8354	10961	4084	654	36
	2004	0	589	7989	11859	4306	716	38
	2005	0	521	8076	13218	4797	793	36
	2006	0	494	7707	13605	5237	839	46
	2007	0	476	7226	13968	5733	849	27
	2008	0	445	6820	14574	6479	883	42
	2009	0	457	6623	14266	6950	950	55
	2010	0	509	6472	14688	7806	1215	59

TABELA A.0.8. *Relação de masculinidade (N.º), grupo etário 0-4 anos, 2001-2006* ^aFonte: INE, *Estimativas Anuais da População Residente*

	Grupo etário
Ano	0-4 anos
2001	105,2
2002	105,9
2003	106,2
2004	106,5
2005	106,9
2006	106,4

^aA relação de masculinidade corresponde ao quociente entre os efetivos populacionais do sexo masculino e os do sexo feminino, habitualmente expresso por 100 mulheres (INE, Sistema de Metainformação).

TABELA A.0.9. *Saldo migratório anual, 2001-2006*Fonte: INE, *Indicadores Demográficos*

Ano	Saldo migratório (HM)
2001	65000
2002	70000
2003	63500
2004	47240
2005	38400
2006	26100

TABELA A.0.10. *Distribuição percentual dos migrantes internacionais por sexo (%), 2000, 2010*

Fonte: Organização das Nações Unidas (UN, 2011)

Grupo etário (anos)	Ano			
	2000		2010	
	Sexo		Sexo	
	H	M	H	M
0-4	51,1	48,9	51,8	48,2
5-9	50,7	49,3	51,4	48,6
10-14	50,2	49,8	51	49
15-19	50,1	49,9	50,8	49,2
20-24	50,2	49,8	50,9	49,1
25-29	50,5	49,5	51,2	48,8
30-34	50,9	49,1	51,6	48,4
35-39	51	49	51,7	48,3
40-44	50,2	49,8	50,9	49,1
45-49	48,3	51,7	49	51
50-54	46,4	53,6	47,1	52,9
55-59	44,8	55,2	45,7	54,3
60-64	41,9	58,1	42,8	57,2
65 +	38,8	61,2	39,7	60,3
Total	49,2	50,8	49,7	50,3

TABELA A.0.11. *Distribuição percentual dos migrantes internacionais por grupo etário e sexo (%), 2000, 2010**Fonte: Organização das Nações Unidas (UN, 2011)*

Sexo	Grupo etário (anos)	Ano	
		2000	2010
H	0-4	1,7	1,9
	5-9	3,7	3,2
	10-14	5,2	4
	15-19	7,5	6,1
	20-24	11,3	9,6
	25-29	18,5	14,4
	30-34	14,5	15,1
	35-39	10,7	12,7
	40-44	8,3	9,8
	45-49	5,8	7,3
	50-54	3,6	4,8
	55-59	2,5	3,4
	60-64	2	2,5
	65 +	4,7	5,2
	Total	100,0	100,0
M	0-4	1,6	1,8
	5-9	3,5	2,9
	10-14	4,9	3,8
	15-19	7,2	5,9
	20-24	10,8	9,1
	25-29	17,6	13,5
	30-34	13,5	14
	35-39	10	11,7
	40-44	8	9,3
	45-49	6,1	7,4
	50-54	4	5,3
	55-59	2,9	4,1
	60-64	2,7	3,3
	65 +	7,1	7,9
	Total	100,0	100,0

TABELA A.0.12. *Taxas de Fecundidade específicas por idade (‰), variante média, 2005-2030*

Fonte: *World Population Prospects: The 2010 Revision (ONU, 2011b)*

Período	Grupo etário (anos)						
	15-19	20-24	25-29	30-34	35-39	40-44	45-49
2005-2010	16,78	45,53	78,49	84,26	39,60	7,61	0,38
2010-2015	12,52	39,29	77,40	85,46	40,34	7,17	0,29
2015-2020	8,82	34,45	78,53	89,07	42,18	6,92	0,23
2020-2025	5,43	31,30	83,90	97,51	46,32	7,06	0,16
2025-2030	1,78	28,08	90,60	107,69	51,31	7,26	0,09

TABELA A.0.13. *Taxas de mortalidade infantil (‰), por sexo, variante média, 2005-2030*

Fonte: *World Population Prospects: The 2010 Revision (ONU, 2011b)*

Período	Sexo		
	HM	H	M
2005-2010	4,5	4,8	4,0
2010-2015	4,2	4,5	3,8
2015-2020	4,0	4,4	3,7
2020-2025	3,9	4,3	3,6
2025-2030	3,8	4,2	3,5

TABELA A.0.14. *Esperança de vida à nascença (anos), por sexo, variante média, 2005-2030*

Fonte: *World Population Prospects: The 2010 Revision (ONU, 2011b)*

Período	Sexo		
	HM	H	M
2005-2010	78,6	75,3	81,8
2010-2015	79,8	76,8	82,8
2015-2020	80,4	77,4	83,3
2020-2025	80,9	78,0	83,8
2025-2030	81,4	78,5	84,3

TABELA A.0.15. *Saldo migratório anual, 2006-2031*

Fonte: INE, Indicadores Demográficos (2006 a 2010) e INE (2008), Cenário Central (2011 a 2031)

Ano	Saldo migratório (HM)	Ano	Saldo migratório (HM)
2006	26100	2019	36584
2007	19500	2020	36584
2008	9361	2021	36584
2009	15408	2022	36584
2010	3815	2023	36584
2011	25712	2024	36584
2012	27265	2025	36584
2013	28818	2026	36584
2014	30371	2027	36584
2015	31925	2028	36584
2016	33477	2029	36584
2017	35030	2030	36584
2018	36584	2031	36584

TABELA A.0.16. *População residente em Portugal, por sexo, grupo etário e nível de escolaridade (2001, 2016, 2031), por cenário*

Fonte: Cálculos do autor

Ano	Grupo etário	H					M					HM				
		SE / 2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	TOTAL	SE / 2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	TOTAL	SE / 2.º Ciclo EB	Ensino Básico	Ensino Secundário	Ensino Superior	TOTAL
2001	0-4	273293	0	0	0	273293	261053	0	0	0	261053	534286	0	0	0	534286
	5-9	272503	0	0	0	272503	259891	0	0	0	259891	532394	0	0	0	532394
	10-14	288882	4558	0	0	293440	276042	4513	0	0	280555	564923	9072	0	0	573995
	15-19	135885	182513	29524	0	347923	93541	195429	45117	0	334087	229427	377942	74641	0	682010
	20-24	127423	123900	130357	14451	396131	90520	96514	168341	31799	387174	217943	220413	298699	46250	783305
	25-29	169220	89870	95316	50819	405224	140507	67083	102837	91230	401656	309727	156952	198152	142049	806880
	30-34	194837	74672	64964	41142	375615	179071	62336	71050	66072	378529	373908	137008	136014	107214	754144
	35-39	217875	66389	56311	34437	375012	215587	59250	60161	53321	388319	433463	125639	116472	87758	763331
	40-44	219754	52816	50002	31423	353995	226450	49251	50123	41711	367535	446204	102067	100125	73134	721530
	45-49	225638	36055	37700	30694	330086	245395	34019	34789	35254	349457	471033	70074	72489	65947	675943
	50-54	219428	32123	31229	23645	306426	248197	29388	27105	25236	329927	467625	61512	58334	48882	636353
	55-59	200293	25506	23777	16665	266241	243895	21297	18238	16294	299724	444188	46803	42014	32959	565965
	60-64	204210	19363	17840	12237	253650	253935	14791	13112	10147	291985	458145	34153	30953	22384	545635
	65-69	203168	15160	14449	9035	241812	260035	12241	12146	6766	291188	463203	27401	26595	15801	533000
	70-74	167331	10213	10288	6841	194673	232088	9078	9446	4335	254947	399419	19291	19734	11176	449620
	75+	232474	11424	11902	8932	264732	399343	12751	12675	5166	429935	631817	24175	24578	14098	694667
	TOTAL	3352155	744564	573658	280320	4950696	3625551	667939	625141	387331	5305962	6977706	1412502	1198799	667651	10256568
2016	0-4	274982	0	0	0	274982	259244	0	0	0	259244	534226	0	0	0	534226
	5-9	290071	0	0	0	290071	273683	0	0	0	273683	563754	0	0	0	563754
	10-14	287516	4541	0	0	292057	268678	4396	0	0	273073	556194	8936	0	0	565130
	15-19	106636	149338	26492	0	282466	69801	161308	39563	0	270671	176437	310646	66055	0	553137
	20-24	106645	50502	118355	10783	286286	69918	31263	143424	29948	274553	176563	81766	261779	40732	560840
	25-29	116249	55646	76364	65271	313530	77037	34823	89724	100969	302554	193285	90470	166088	166241	616083
	30-34	143276	67417	88226	74866	373785	99879	42510	104454	115704	362547	243155	109927	192680	190570	736332
	35-39	140754	130052	83640	73230	427677	103004	102059	103927	113081	422072	243759	232111	187567	186312	849749
	40-44	181173	93311	98433	53690	426606	153230	71016	107220	96221	427688	334403	164326	205653	149912	854294
	45-49	200921	75971	66139	42650	385681	188691	64913	73222	68888	395714	389613	140884	139361	111538	781396
	50-54	215561	65537	56009	35089	372196	220544	60445	61017	54775	396781	436106	125982	117026	89864	768977
	55-59	207569	50135	48144	31313	337161	226058	48949	49779	42372	367157	433626	99084	97923	73685	704318
	60-64	200191	32629	34901	29605	297327	236805	32968	33838	35115	338726	436996	65597	68739	64721	636053
	65-69	178505	28957	27365	22041	254868	229345	27365	25642	24676	307028	407849	54322	53007	46717	561896
	70-74	142601	19095	19155	14772	159623	210530	18683	16538	15531	261282	353131	37778	35693	30303	456906
	75+	232426	19885	23013	19589	294913	438235	21570	23236	15910	408950	670661	41455	46249	35499	793864
	TOTAL	3025076	841015	764637	472901	5105229	3124682	722269	871583	713191	5431726	6149758	1563284	1637820	1186092	10536584
Cenário constante	0-4	258867	0	0	0	258867	243979	0	0	0	243979	502846	0	0	0	502846
	5-9	256872	0	0	0	256872	242112	0	0	0	242112	498984	0	0	0	498984
	10-14	261704	4132	0	0	265836	246737	4036	0	0	250772	508440	8168	0	0	516609
	15-19	107934	151092	26816	0	285842	69685	160984	39502	0	270172	177620	312076	66318	0	556013
	20-24	113815	54248	126454	11554	306071	73709	33283	151562	31702	290256	187523	87531	278017	43256	596327
	25-29	115828	57496	77251	65922	316497	74982	35404	88424	99567	298377	190810	92900	165674	165489	614874
	30-34	115322	57991	76298	64894	314504	76830	36896	89178	100348	303252	192152	94887	165476	165242	617757
	35-39	119070	58983	77131	65414	320598	80349	37917	88993	101052	309212	199419	96901	167024	166466	629810
	40-44	129998	61827	81317	68253	342395	89274	40265	95602	107543	332684	219272	102092	176918	176796	675078
	45-49	153845	70808	90530	76991	392174	111377	46549	107654	119455	385035	265222	117357	198184	196446	777209
	50-54	146325	128675	83554	73837	432391	112236	103565	104951	114817	435569	258561	232240	188505	188554	867960
	55-59	177244	89958	95415	53452	416068	157766	71217	106355	96628	431966	335009	161175	201770	150080	848035
	60-64	185053	70149	62218	41595	359015	186993	63295	71527	68507	390321	372046	133444	133744	110102	749336
	65-69	182328	56731	50168	33129	323266	209186	56965	58178	53586	377915	392424	113696	108346	86715	701181
	70-74	156466	39367	40026	28158	264016	201418	43851	45705	40429	331402	357884	83217	85730	68587	595418
	75+	244946	38411	47240	47122	377718	455301	48514	52728	55879	612422	700247	86925	99967	103000	990140
	TOTAL	2726527	939687	934416	631321	5323131	2631933	782741	1101259	899513	5505447	5358461	1722609	2035674	1620834	10737578
2016	0-4	272798	0	0	0	272798	257182	0	0	0	257182	529980	0	0	0	529980
	5-9	289372	0	0	0	289372	273022	0	0	0	273022	562394	0	0	0	562394
	10-14	287516	4541	0	0	292057	268678	4396	0	0	273073	556194	8936	0	0	565130
	15-19	69424	162789	50254	0	282466	38849	170925	60897	0	270671	108273	333714	111150	0	553137
	20-24	84047	35953	139735	26545	286280	50837	35960	136019	48145	274562	134885	75513	275754	74690	560842
	25-29	109682	34728	102289	67070	313770	71384	32285	70708	128190	302566	181066	67012	172997	195260	616336
	30-34	143276	58217	89360	82978	373832	99879	41813	83717	137153	362562	243155	100030	173077	220131	736394
	35-39	140754	130052	88297	68557	427660	103004	102059	97815	119201	422080	243759	232111	186112	187759	849741
	40-44	181173	93311	98433	53690	426606	153230	71016	107220	96221	427688	334403	164326	205653	149912	854294
	45-49	200921	75971	66139	42650	385681	188691	64913	73222	68888	395714	389613	140884	139361	111538	781396
	50-54	215561	65537	56009	35089	372196	220544	60445	61017	54775	396781	436106	125982	117026	89864	768977
	55-59	207569	50135	48144	31313	337161	226058	48949	49779	42372	367157	433626	99084	97923	73685	704318
	60-64	200191	32629	34901	29605	297327	236805	32968	33838	35115	338726	436996	65597	68739	64721	636053
	65-69	178505	28957	27365	22041	254868	229345	27365								

APÊNDICE B

Gráficos auxiliares

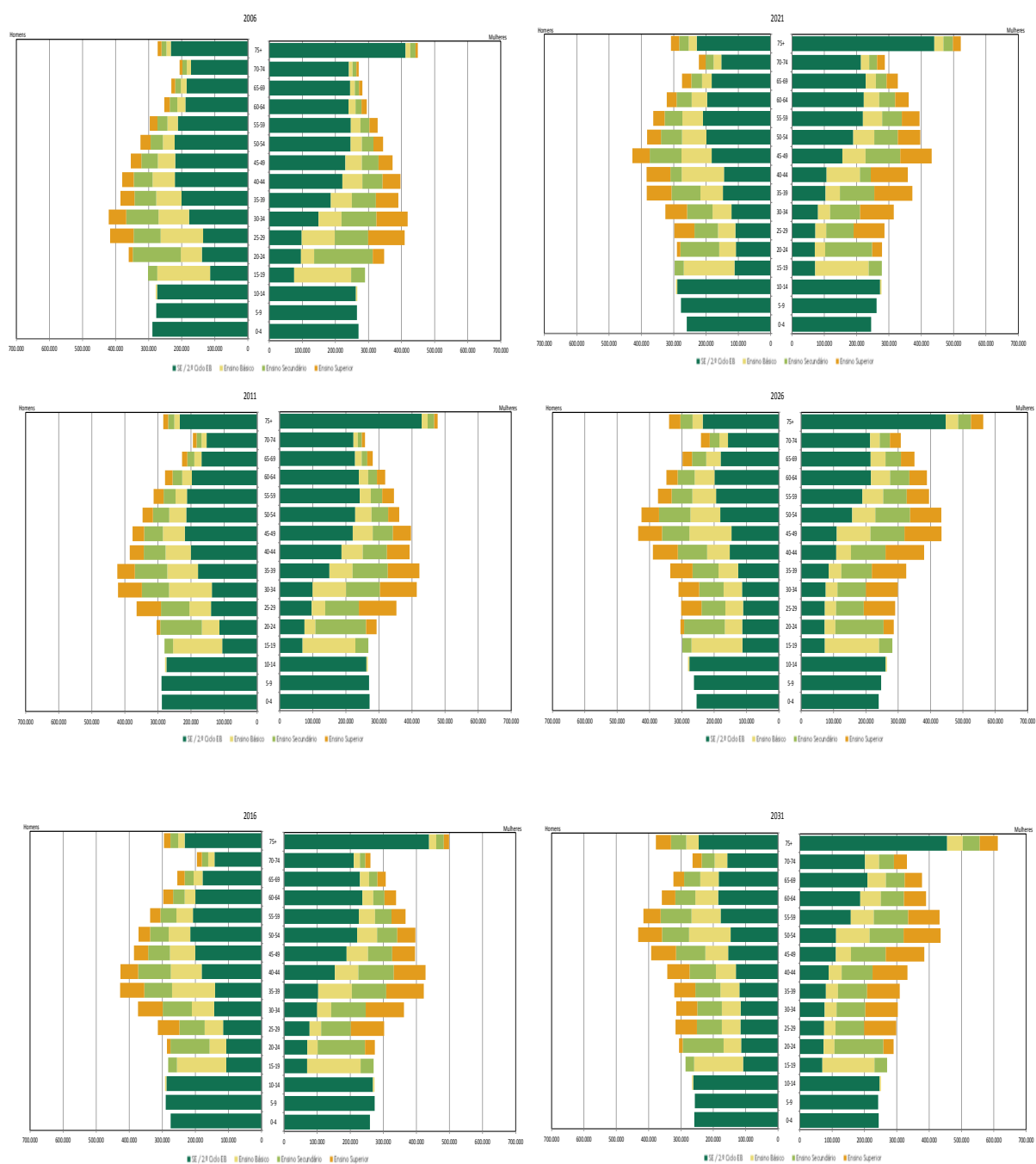


FIGURA B.0.1. *População residente em Portugal, por sexo, grupo etário e nível de escolaridade, 2001-2031, Cenário Constante*

Fonte: Cálculos do autor

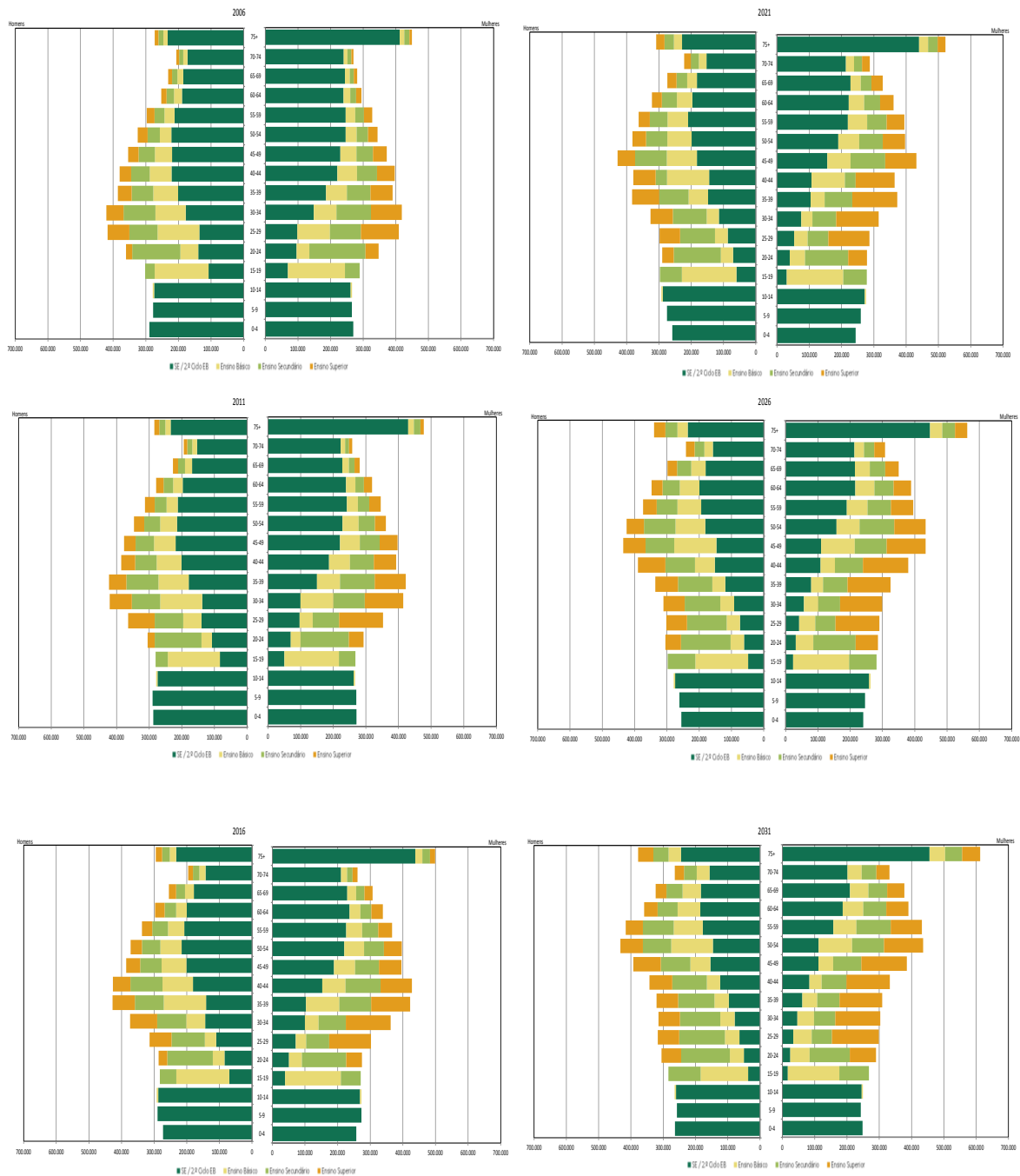


FIGURA B.0.2. *População residente em Portugal, por sexo, grupo etário e nível de escolaridade, 2001-2031, Cenário Tendência*

Fonte: Cálculos do autor

APÊNDICE C

Expressão matricial para projeção da população feminina para o quinquênio 2001-2006

Com o intuito de ilustrar os cálculos realizados, apresenta-se abaixo, e a título exemplificativo, a expressão matricial utilizada para a projeção da população feminina para o dia 1 de janeiro de 2006, partindo da população inicial (1 de janeiro de 2001) e em que se utilizam os parâmetros calculados nas secções 5.2.3.1, 5.2.3.2, 5.2.3.3 e 5.2.3.4:

$$\mathbf{B}\ell_{2001} + \mathbf{SM}_{2001-2006} = \ell_{2006} \Leftrightarrow$$

$$\begin{bmatrix}
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,092 & 0,029 & 0,040 & 0,000 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0,998 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0,983 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0,016 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,262 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,595 & 0,613 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,142 & 0,386 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,998 & 0 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,184 & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,815 & 0,254 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,744 & 0 & 0 & \dots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0,770
 \end{bmatrix}
 \begin{bmatrix}
 261053 \\
 0 \\
 0 \\
 0 \\
 0 \\
 259891 \\
 0 \\
 0 \\
 0 \\
 0 \\
 276042 \\
 4513 \\
 0 \\
 0 \\
 93541 \\
 195429 \\
 45117 \\
 0 \\
 \vdots \\
 5166
 \end{bmatrix}
 +
 \begin{bmatrix}
 2121 \\
 0 \\
 0 \\
 0 \\
 0 \\
 4527 \\
 0 \\
 0 \\
 0 \\
 0 \\
 6208 \\
 102 \\
 0 \\
 0 \\
 2534 \\
 5460 \\
 1312 \\
 0 \\
 \vdots \\
 67
 \end{bmatrix}
 =
 \begin{bmatrix}
 269621 \\
 0 \\
 0 \\
 0 \\
 0 \\
 265128 \\
 0 \\
 0 \\
 0 \\
 0 \\
 261655 \\
 4282 \\
 0 \\
 0 \\
 74877 \\
 172425 \\
 42178 \\
 0 \\
 \vdots \\
 7385
 \end{bmatrix}$$

Exemplo de tratamento de dados agregados no contexto de Cadeias de Markov

Kalbfleisch e Lawless (1984) apresentaram uma forma de estimar as probabilidades de transição de uma cadeia de Markov homogénea no caso em que se dispõe unicamente de dados agregados. Admitimos que as observações são feitas sobre um grupo de indivíduos, de dimensão n , que agem independentemente uns dos outros, e que cada um deles atravessa os estados de acordo com um processo de Markov homogéneo de espaço de estados $E = 1, 2, \dots, k$ e matriz $k \times k$ de intensidades de transição $\mathbf{Q} = \{q_{ij}(\theta)\}$; as intensidades de transição são funções diferenciáveis do vetor de parâmetros $\Theta = (\theta_1, \dots, \theta_r)'$. Supomos ainda que as observações são feitas em instantes $t = 0, 1, 2, \dots, m$, e que os dados disponíveis consistem somente no número total de indivíduos, N_{it} , no estado i no instante t , $i = 1, \dots, k$ e $t = 0, 1, \dots, m$. Admitimos, para já, que o sistema é fechado, ou seja, que o mesmo número de indivíduos, $n = \sum_{i=1}^k N_{it}$, é observado em cada instante t , e que o número de indivíduos em cada um dos estados no instante $t = 0$ é conhecido. Podemos definir os vectores:

$$\mathbf{W}'_t = n^{-1}\{N_{1t}, \dots, N_{kt}\}$$

e

$$\mathbf{Z}'_t = n^{-1}\{N_{1t}, \dots, N_{k-1,t}\}$$

para $t = 0, 1, \dots, m$. Seja ainda Y_{ijt} o número de indivíduos que ocupam o estado i no instante $t - 1$ e o estado j no instante t ; contudo, encontramos-nos no caso em que apenas os valores de N_{it} são observados, e não os de Y_{ijt} .

Sendo $\mathbf{N}'_t = \{N_{1t}, \dots, N_{kt}\}$, consideremos a distribuição conjunta de $\mathbf{N}_1, \dots, \mathbf{N}_m$, dado \mathbf{N}_0 . Em particular, dado $N_{i,t-1}$, $\mathbf{Y}_{it} = \{Y_{i1t}, \dots, Y_{ikt}\}$ tem uma distribuição multinomial de k classes com parâmetros $\{N_{i,t-1}; p_{i1}, \dots, p_{ik}\}$ onde a matriz $\mathbf{P} = (p_{ij})$ é a matriz de probabilidades de transição. Dado que

$$N_{it} = Y_{1it} + \dots + Y_{kit} \quad i = 1, \dots, k, \quad t = 1, \dots, m \quad (5)$$

ou seja, em termos matriciais:

$$n\mathbf{W}_t = \sum_{i=1}^k \mathbf{Y}_{it} \quad t = 1, \dots, m$$

a distribuição condicional de \mathbf{W}_t , dado \mathbf{W}_{t-1} , é uma convolução de multinomiais. Como $\{\mathbf{W}_t : t = 0, 1, 2, \dots, m\}$ é um processo de Markov, a função de probabilidade conjunta de $\mathbf{W}_1, \dots, \mathbf{W}_m$,

dado \mathbf{W}_0 , é construída como o produto de probabilidades condicionais na forma:

$$P(\mathbf{W}_1|\mathbf{W}_0)P(\mathbf{W}_2|\mathbf{W}_1)\dots P(\mathbf{W}_m|\mathbf{W}_{m-1}) = \prod_{t=1}^m P(\mathbf{W}_t|\mathbf{W}_{t-1})$$

A construção da função de verosimilhança requer a obtenção da função de probabilidade de \mathbf{W}_t , o que torna a estimação direta de \mathbf{P} por máxima verosimilhança computacionalmente intratável. No entanto, os momentos condicionais (e funções geradoras) de \mathbf{W}_t , dado \mathbf{W}_{t-1} , são facilmente obtidos, e podem ser utilizados para desenvolver procedimentos de estimação baseados em mínimos quadrados ou máxima verosimilhança aproximada. A partir da equação 5 verifica-se que:

$$E[\mathbf{W}_t|\mathbf{W}_{t-1}] = \mathbf{P}'\mathbf{W}_{t-1}$$

e

$$nCov(\mathbf{W}_t|\mathbf{W}_{t-1}) = \Sigma_t = diag\{\mathbf{P}'\mathbf{W}_{t-1}\} - \mathbf{P}'diag\{\mathbf{W}_{t-1}\}\mathbf{P}$$

para $t = 0, 1, 2, \dots, m$. Para além disso, a distribuição condicional de \mathbf{W}_t , dado \mathbf{W}_{t-1} , é singular, aproximadamente normal k -variada e, como tal, é mais conveniente trabalhar com a distribuição não singular equivalente de \mathbf{Z}_t , dado \mathbf{W}_{t-1} , aproximadamente normal $(k-1)$ -variada com:

$$E[\mathbf{Z}_t|\mathbf{W}_{t-1}] = \mathbf{P}_1'\mathbf{W}_{t-1} \quad (6)$$

e

$$nCov(\mathbf{Z}_t|\mathbf{W}_{t-1}) = \Sigma_{1t} = diag\{\mathbf{P}_1'\mathbf{W}_{t-1}\} - \mathbf{P}_1'diag\{\mathbf{W}_{t-1}\}\mathbf{P}_1$$

para $t = 0, 1, 2, \dots, m$, em que \mathbf{P}_1 é a matriz $k \times (k-1)$ obtida pela eliminação da última coluna de \mathbf{P} e Σ_{1t} é a submatriz principal de Σ_t , de dimensão $(k-1) \times (k-1)$.

Para estimar a matriz de probabilidades de transição \mathbf{P} , podem ser consideradas estimativas de mínimos quadrados condicionais, obtidas pela minimização de formas quadráticas do tipo:

$$S_Q = n \sum_{t=1}^m (\mathbf{Z}_t - \mathbf{P}_1'\mathbf{W}_{t-1})' \mathbf{Q}_{t-1} (\mathbf{Z}_t - \mathbf{P}_1'\mathbf{W}_{t-1})$$

baseadas na expressão 6, em que $\mathbf{Q}_0, \mathbf{Q}_1, \dots, \mathbf{Q}_{m-1}$ são matrizes definidas positivas de dimensão $(k-1) \times (k-1)$ adequadamente escolhidas. Os elementos de \mathbf{Q}_{t-1} podem depender de \mathbf{W}_t , mas são funcionalmente independentes das probabilidades de transição p_{ij} .

Suponhamos que \mathbf{P} contém $r \leq k(k-1)$ probabilidades de transição linearmente independentes p_{ij} a serem estimadas. Notemos que $r < k(k-1)$ é possível se \mathbf{P} contém zeros estruturais, ou seja, elementos que se sabe serem nulos *a priori*. Escrevemos as probabilidades a serem estimadas como um vetor coluna $\boldsymbol{\gamma}$, de dimensão $r \times 1$, pela ordenação resultante de percorrermos \mathbf{P}_1 coluna a coluna. Definimos também matrizes \mathbf{B}_t tais que:

$$\mathbf{B}_t\boldsymbol{\gamma} = \mathbf{P}_1'\mathbf{W}_t \quad t = 0, 1, 2, \dots, m$$

Desta forma, S_Q pode ser reescrito como:

$$S_Q = n \sum_{t=1}^m (\mathbf{Z}_t - \mathbf{B}_{t-1}\gamma)' \mathbf{Q}_{t-1} (\mathbf{Z}_t - \mathbf{B}_{t-1}\gamma)$$

e, no geral, desde que $\sum_{t=1}^m \mathbf{B}_{t-1}' \mathbf{B}_{t-1}$ seja não singular (ou seja, invertível e, portanto, todos os parâmetros de γ sejam estimáveis), S_Q é minimizado por:

$$\tilde{\gamma}_Q = \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \mathbf{Q}_{t-1} \mathbf{B}_{t-1} \right)^{-1} \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \mathbf{Q}_{t-1} \mathbf{Z}_t \right)$$

Podem ser considerados dois estimadores para γ :

- (1) O estimador de mínimos quadrados ordinários, quando $\mathbf{Q}_{t-1} = \mathbf{I}_{k-1}$ onde \mathbf{I}_{k-1} é a matriz identidade de ordem $k-1$. Neste caso, teremos:

$$\tilde{\gamma}_{OLS} = \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \mathbf{B}_{t-1} \right)^{-1} \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \mathbf{Z}_t \right)$$

- (2) O estimador de mínimos quadrados ponderados, em que \mathbf{Q}_{t-1} é substituído por uma estimativa de $\Sigma_{t-1}^{-1}(\gamma)$. O estimador será dado por:

$$\tilde{\gamma}_{WLS} = \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \tilde{\Sigma}_{t-1}^{-1}(\gamma) \mathbf{B}_{t-1} \right)^{-1} \left(\sum_{t=1}^m \mathbf{B}_{t-1}' \tilde{\Sigma}_{t-1}^{-1}(\gamma) \mathbf{Z}_t \right)$$

Podemos partir de uma estimativa para γ que resulte de considerar $\mathbf{Q}_{t-1} = \hat{\Sigma}_{t-1}^{-1} = \Sigma_{t-1}^{-1}(\tilde{\gamma}_{OLS})$ e conduzir um processo iterativo através da estimação de $\Sigma_{t-1}^{-1}(\gamma)$ em cada passo, usando $\tilde{\gamma}_{WLS}$ da iteração anterior.

Devemos notar que algumas das estimativas de p_{ij} podem situar-se fora do intervalo $(0, 1)$ e, por isso, a estimação deve ser sujeita a restrições.

Considerámos, até agora, que o sistema em estudo é fechado, ou seja, que o número de indivíduos em observação se mantém constante ao longo do tempo. Sabemos, no entanto, que, na realidade, a população em estudo está sujeita à entrada de novos elementos, seja através da imigração (que pode ocorrer em qualquer grupo etário) como dos nascimentos, designados por entradas endógenas e que ocorrem somente no primeiro grupo etário, bem como à saída de indivíduos por morte ou emigração. No que se refere à entrada de elementos, devemos, antes de mais, assumir que, uma vez no sistema, os novos elementos agem independentemente dos restantes e movem-se de acordo com as mesmas leis de probabilidade.

Para tratar a possibilidade de saída da população por morte, podemos considerar a existência de um estado adicional. No instante inicial, t_0 , não existem indivíduos neste estado e, à medida que o tempo decorre, todos aqueles que morrem são alocados, cumulativamente, a este estado. Desta forma, estudamos as probabilidades de morte a partir de cada estado como as probabilidades de transição

para este estado específico. Este procedimento não exige que conheçamos o estado de origem dos indivíduos que falecem em cada intervalo de tempo, informação esta que é muitas vezes difícil de obter.

O tratamento das migrações pode basear-se nos dados relativos ao saldo migratório, resultando em informação sobre o balanço conjunto entre as imigrações e emigrações esperado em cada intervalo. Admitimos que as entradas e saídas por migração ocorrem imediatamente após cada observação, e reinterpretamos $N_{i,t-1}$ como o número de indivíduos no estado i quando as migrações estão completas após a observação no instante $t-1$. O número de indivíduos no i -ésimo estado após a observação no instante t , mas antes de migrarem novos elementos, é dado por M_{it} . Assim, $n_{t-1} = \sum_{i=1}^k N_{i,t-1}$ representa o número total de indivíduos “em risco” na observação em t , e definimos $\mathbf{W}'_t = n_{t-1}^{-1} \{N_{1t}, \dots, N_{kt}\}$, e $\mathbf{Z}'_t = n_{t-1}^{-1} \{M_{1t}, \dots, M_{k-1,t}\}$. Todos os resultados anteriores aplicam-se agora com pequenas modificações.

